

Trabajo Fin de Grado

“Estudio estadístico de la evolución e
impacto de la crisis económica en el
sector de la construcción, sector
inmobiliario, e industria metalúrgica
en España mediante modelos ARIMA
y detección de outliers”

Froilán Martínez Martínez De Ibarreta

Tutora

Teresa Villagarcía Casla

Leganés – septiembre / octubre de 2018

ÍNDICE DE CONTENIDOS.

1	Introducción y objetivos.	12
2	Estructura económica española. Evolución del sector de la construcción, sector inmobiliario y de la industria metalúrgica.	14
2.1	Estructura de la economía española.	14
2.2	Evolución histórica del sector de la construcción.	17
2.3	Evolución del sector inmobiliario.	19
2.4	Industria metalúrgica.	21
3	Introducción teórica al análisis de series temporales.	22
3.1	Evolución histórica del análisis de series temporales.	22
3.2	Concepto de serie temporal.	23
3.3	Características de una serie temporal. Definición de serie estacionaria.	23
3.4	Estructuras de dependencia: función de autocorrelación simple (FAS) y función de autocorrelación parcial (FAP).	24
3.5	Ruido Blanco.	26
3.6	Operadores matemáticos utilizados en el análisis de series temporales.	27
3.7	Modelos autorregresivos, AR(p).	28
3.8	Modelos de media móvil, MA (q).	29
3.9	Modelos ARMA (p, q).	31
3.10	Estacionalidad.	32
3.11	Modelos integrados ARIMA (p, d, q) x (P, D, Q) s.	33
3.12	Intervenciones.	33
3.13	Datos atípicos: outliers.	36
3.14	Metodología de Box-Jenkins para el análisis de series temporales.	40
4	Series analizadas.	41
4.1	Trabajadores asociados a la seguridad social en el sector de la construcción.	43
4.2	Índice nacional del precio de la mano de obra en el sector de la construcción en España.	47
4.3	Parados en el sector de la construcción.	52

4.4	Número de asalariados en el sector de la construcción.	61
4.5	Número de operaciones de compraventa de vivienda nueva.....	63
4.6	Número de operaciones de compraventa de vivienda de segunda mano.....	66
4.7	Hipotecas totales constituidas en España.	69
4.8	Índice de producción industrial (IPI) de la industria metalúrgica en España.	73
4.9	Producción interna de acero en España.	76
4.10	Exportaciones de acero.....	79
4.11	Importaciones de acero en España.	82
4.12	Paro total registrado en España.	86
4.13	IPC: índice de precios de consumo.	89
5	Análisis de resultados.....	93
5.1	Impacto de la crisis en el sector de la construcción, el sector inmobiliario y la industria metalúrgica.	93
5.1.1	Tabla de outliers detectados en todas las series.....	93
5.1.2	Impacto y efectos de la crisis sobre los sectores estudiados.....	97
5.1.3	Posibles explicaciones y causas para los atípicos detectados.....	99
5.2	Predicciones con modelos ARIMA.	101
5.2.1.	Diciembre de 2002.	102
5.2.3.	Abril de 2010.....	109
5.2.4.	Septiembre de 2017	113
5.2.5.	Conclusiones sobre las predicciones con modelos ARIMA.....	116
6	Conlusiones	117
7	Anexo: COdigo en r (V 3.5.1) para el análisis de series temporales.	118
8	Bibliografía	122

ÍNDICE DE FIGURAS.

Figura 2.1. Evolución del peso porcentual de los distintos sectores en España sobre el PIB desde el año 1980. Fuente: Elaboración propia a partir de datos extraídos del INE.	14
Figura 2.2. Evolución de la estructura porcentual del empleo en España desde el año 1980 (en porcentaje). Fuente: Elaboración propia a partir de datos extraídos del INE.	15
Figura 2.3. Inversión en I+D en la industria española en comparación con UE15 en el año 2010 (en porcentaje). Fuente: Eurostat.	16
Figura 2.4. Crecimiento del Índice de Producción Industrial, desde 1993 hasta 2018 (en porcentaje). Fuente: Elaboración propia a partir de datos extraídos del INE.	17
Figura 2.5. Evolución del número de empleados del sector de la construcción desde 1995 hasta 2016 (en miles). Fuente: Elaboración propia a partir de datos extraídos del INE.	18
Figura 2.6. Consumo de cemento en España desde 1996 hasta 2016 (en miles de toneladas). Fuente: Elaboración propia a partir de datos extraídos del INE.	19
Figura 2.7. Evolución del precio de la vivienda en España desde 1985 hasta 2012 (precio medio del m2 en euros). Fuente: INE.	20
Figura 2.8. Oferta y demanda de vivienda en España desde el 2006 hasta el 2015. Fuente: INE y Banco de España realizada por Bankinter.	21
Figura 3.1. Evolución histórica del análisis de series temporales.	23
Figura 3.2. Posibles formas de influencia de las observaciones sobre las demás.	25
Figura 3.3. Gráficas FAS y FAP de una serie temporal.	26
Figura 3.4. Gráficas FAS y FAP de un ruido blanco.	26
Figura 3.5. FAS y FAP para modelos AR (1). (Peña 2010).	29
Figura 3.6. FAS y FAP para algunos modelos AR (2). (Peña 2010).	29
Figura 3.7. FAS y FAP para modelos MA (1). (Peña 2010).	30
Figura 3.8. FAS y FAP para algunos modelos MA (q). (Peña 2010).	31
Figura 3.9. FAS y FAP de modelos ARMA. (Peña 2010).	32
Figura 3.10. Efecto de un AO sobre una serie temporal (elaboración en R).	37
Figura 3.11. Efecto de un IO sobre una serie temporal (elaboración en R).	38
Figura 3.12. Efecto de un LS sobre una serie temporal (elaboración en R).	39
Figura 3.13. Efecto de un cambio transitorio sobre una serie temporal (elaboración en R).	40
Figura 4.1. Gráfico de secuencia para la serie temporal número de asociados a la seguridad social en el sector de la construcción.	43
Figura 4.2. FAS y FAP del número de asociados a la seguridad social en el sector de la construcción.	44
Figura 4.3. Gráfico de secuencia, FAS y FAP de la serie para el número de asociados a la seguridad social en el sector de la construcción transformada en estacionaria.	44

Figura 4.4. FAS y FAP residual para el ajuste de la serie número de asociados a la seguridad social en el sector de la construcción.....	45
Figura 4.5. Impacto de los outliers en la serie temporal número de asociados a la seguridad social en el sector de la construcción.....	46
Figura 4.6. Gráfico de secuencia para la serie de índice de precios de la mano de obra en el sector de la construcción.....	47
Figura 4.7. FAS para la serie del índice de precios de la mano de obra en el sector de la construcción.	48
Figura 4.8. FAP para la serie del índice de precios de la mano de obra en el sector de la construcción.	48
Figura 4.9. Serie del índice de precios de la mano de obra en el sector de la construcción transformada en estacionaria.	49
Figura 4.10. FAS del índice de precios de la mano de obra en el sector de la construcción transformada en estacionaria.	49
Figura 4.11. FAP del índice de precios de la mano de obra en el sector de la construcción transformada en estacionaria.	50
Figura 4.12. FAS de los residuos para el índice de precios de la mano de obra en el sector de la construcción.	50
Figura 4.13. FAP de los residuos para el índice de precios de la mano de obra en el sector de la construcción.	51
Figura 4.14. Gráfica de la serie índice de precios de la mano de obra en el sector de la construcción con los efectos de los outliers detectados.	52
Figura 4.15. Gráfico de secuencia de los parados en el sector e la construcción en España.	53
Figura 4.16. Descomposición de la serie temporal "parados en el sector de la construcción" en series aditivas de tendencia, estacionalidad y valores aleatorios.....	54
Figura 4.17. FAS parados en el sector de la construcción.....	54
Figura 4.18. FAP parados en el sector de la construcción.....	55
Figura 4.19. Serie temporal "parados en el sector de la construcción" tras tomar una diferencia estacional.	56
Figura 4.20. Serie temporal "parados en el sector de la construcción" con una diferencia regular y una diferencia estacional.	56
Figura 4.21. Serie de los parados en el sector de la construcción transformada en estacionaria tras aplicar una diferencia estacional, una diferencia estacional y tomar el logaritmo natural de la serie.	57
Figura 4.22. FAS parados en el sector de la construcción (estacionaria).	57
Figura 4.23. FAP parados en el sector de la construcción (estacionaria).	58

Figura 4.24. Funciones de autocorrelación para el modelo ARIMA propuesto para la serie de parados en el sector de la construcción.	58
Figura 4.25. Gráfica de la serie "parados del sector de la construcción" con los efectos de los outliers detectados.	60
Figura 4.26. Gráfico de secuencia para la serie trimestral asalariados en el sector de la construcción.	61
Figura 4.27. FAS para la serie trimestral asalariados en el sector de la construcción.	61
Figura 4.28. FAP de secuencia para la serie trimestral asalariados en el sector de la construcción.	62
Figura 4.29. FAS y FAP de los residuos del modelo ARIMA ajustado para la serie asalariados en el sector de la construcción.	63
Figura 4.30. Gráfico de secuencia del número de operaciones de compraventa de viviendas nuevas.	63
Figura 4.31. FAS y FAP para la serie "número de operaciones de compraventa de viviendas nuevas".	64
Figura 4.32. Serie de las operaciones de compraventa de vivienda nueva transformada en estacionaria junto a su FAS y FAP.	64
Figura 4.33. FAS y FAP de los residuos generados por el modelo.	65
Figura 4.34. Efecto de los atípicos sobre la serie número de operaciones de compraventa de vivienda nueva.	66
Figura 4.35. Gráfico de secuencia de la serie operaciones de compraventa de viviendas de 2ª mano.	66
Figura 4.36. FAS y FAP para el número de operaciones de compraventa de viviendas de segunda mano.	67
Figura 4.37. Serie del número de operaciones de compraventa de viviendas de segunda mano transformada en estacionaria, junto a su FAS y FAP.	67
Figura 4.38. FAS y FAP de los residuos generados para el modelo.	68
Figura 4.39. Efecto de los outliers sobre la serie del número de operaciones de compraventa de viviendas de segunda mano.	69
Figura 4.40. Gráfico de secuencia de la serie temporal hipotecas totales constituidas en España. ..	69
Figura 4.41. Funciones de autocorrelación para la serie de hipotecas constituidas en España.	70
Figura 4.42. Serie hipotecas totales constituidas en España transformada en estacionaria junto a sus funciones de autocorrelación FAS y FAP.	70
Figura 4.43. FAS y FAP de los residuos del modelo ARIMA para hipotecas totales constituidas en España.	71
Figura 4.44. Efectos de los outliers sobre el total de las hipotecas constituidas en España.	72
Figura 4.45. Gráfico de secuencia del IPI en la industria metalúrgica.	73
Figura 4.46. FAS y FAP de la serie del IPI en la industria metalúrgica.	73

Figura 4.47. Serie IPI de la industria metalúrgica transformada en estacionaria.	74
Figura 4.48. Residuos del modelo ajustado para IPI en la industria metalúrgica.	75
Figura 4.49. Efectos de los atípicos sobre la serie IPI en la industria metalúrgica.	75
Figura 4.50. Gráfico de secuencia de la serie temporal producción interna de acero.	76
Figura 4.51. FAS y FAP junto al gráfico de la serie producción interna de acero.	76
Figura 4.52. Serie producción interna de acero tras tomar una diferencia regular y otra estacional, junto a su FAS y FAP.	77
Figura 4.53. Residuos generados para el ajuste del modelo ARIMA propuesto para la serie producción interna de acero.	78
Figura 4.54. Efecto de los atípicos en la serie producción interna de acero.	78
Figura 4.55. Gráfico de secuencia de la serie exportación de acero.	79
Figura 4.56. Funciones de autocorrelación para la serie exportaciones de acero.	79
Figura 4.57. Serie exportaciones de acero transformada en estacionara junto a su FAS y FAP.	80
Figura 4.58. Residuos generados en el ajuste para la serie exportaciones de acero junto a su FAS y FAP.	81
Figura 4.59. Efecto de los atípicos sobre la serie exportaciones de acero.	82
Figura 4.60. Gráfico de secuencia para la serie importaciones de acero.	83
Figura 4.61. FAS y FAP de la serie importaciones de acero junto a su gráfico de secuencia.	83
Figura 4.62. Serie importaciones de acero transformada en estacionaria.	84
Figura 4.63. FAS y FAP de los residuos del modelo ajustado para la serie importaciones de acero.	85
Figura 4.64. Efecto de los atípicos sobre las series de las importaciones de acero.	86
Figura 4.65. Gráfico de secuencia de la serie temporal del paro total registrado.	86
Figura 4.66. FAS y FAP de la serie del paro total registrado, junto a su gráfico de secuencia.	87
Figura 4.67. Serie del paro total registrado transformada en estacionaria, junto a su FAS y FAP. .	87
Figura 4.68. Residuos generados por el modelo para la serie del paro total registrado.	88
Figura 4.69. Efecto de los atípicos sobre la serie del paro total registrado.	89
Figura 4.70. Gráfico de secuencia para la serie temporal del IPC nacional.	89
Figura 4.71. Gráfico de secuencia de la serie del IPC, junto a su FAS y FAP.	90
Figura 4.72. Serie del IPC transformada en estacionaria.	90
Figura 4.73. Residuos generados en el ajuste para la serie del IPC.	91
Figura 4.74. Efecto de los atípicos sobre la serie del IPC.	92
Figura 5.1. Evolución anual del número total de atípicos detectados.	94
Figura 5.2. Atípicos totales detectados.	95
Figura 5.3. Distribución de los atípicos detectados en el año 2008.	96
Figura 5.4. Tipos de atípicos detectados en 2008.	97

Figura 5.5. Filtración de la crisis a través de los sectores estudiados, reflejado en los atípicos detectados.	98
Figura 5.6. Serie de los asociados a la seguridad social en el sector de la construcción con los 3 cortes indicados.	102
Figura 5.7. Serie de los asociados a la seguridad social en el sector de la construcción desde enero de 1996 hasta diciembre de 2002.	102
Figura 5.8. Gráfico de secuencia para las predicciones a partir de diciembre de 2002, junto a sus límites de confianza superior e inferior.	104
Figura 5.9. Gráfico de secuencia para las predicciones a partir de diciembre de 2002 en contraposición a los datos reales.	104
Figura 5.10. Serie de asociados a lo seguridad social en el sector de la construcción desde enero de 1996 hasta marzo de 2007.	106
Figura 5.11. Gráfico de secuencia para las predicciones a partir de marzo de 2007, junto a sus límites de confianza superior e inferior.	108
Figura 5.12. Gráfico de secuencia para las predicciones a partir de marzo de 2007 en contraposición a los datos reales.	108
Figura 5.13. Serie del número de asociados a la seguridad social desde enero de 1996 hasta abril de 2010.	109
Figura 5.14. Gráfico de secuencia para las predicciones a partir de abril de 2010, junto a sus límites de confianza superior e inferior.	111
Figura 5.15. Gráfico de secuencia para las predicciones a partir de abril de 2010, junto a sus límites de confianza superior e inferior, y contrapuestos con os datos reales.	112
Figura 5.16. Serie de asociados a la seguridad social en el sector de la construcción.	113
Figura 5.17. Predicciones para la serie de los asociados a la seguridad social.	114
Figura 5.18. Predicciones para la serie de la seguridad social comparados con los nuevos datos publicados.	114

ÍNDICE DE TABLAS.

Tabla 4.1. Listado de series temporales a analizar.	43
Tabla 4.2. Atípicos detectados en SPSS para la serie temporal “número de asociados a la seguridad social en el sector de la construcción”.....	47
Tabla 4.3. Valores atípicos detectados para la serie de índice de precios de mano de obra en el sector de la construcción.	52
Tabla 4.4. Parámetros modelo ARIMA para la serie de parados en el sector de la construcción. ...	59
Tabla 4.5. Resultados generales para el ajuste del modelo ARIMA propuesto para la serie temporal de parados en el sector de la construcción.....	59
Tabla 4.6. Valores atípicos detectados para la serie de parados en el sector de la construcción.....	60
Tabla 4.7. Atípicos detectados para la serie del número de operaciones compraventa de viviendas nuevas.....	65
Tabla 4.8. Atípicos para la serie operaciones de compraventa de vivienda de 2ª mano.....	68
Tabla 4.9. Atípicos detectados en R para la serie temporal total de hipotecas constituidas en España.	72
Tabla 4.10. Atípicos detectados en R para la serie IPI en la industria metalúrgica.....	75
Tabla 4.11. Atípicos detectados para la serie producción interna de acero.	78
Tabla 4.12. Atípicos detectados para la serie exportaciones de acero.....	81
Tabla 4.13. Atípicos detectados para la serie importaciones de acero.	85
Tabla 4.14. Atípicos detectados para la serie del paro total.	88
Tabla 4.15. Atípicos detectados para la serie del IPC.	91
Tabla 5.1. Outliers detectados en el conjunto total de series analizadas.	94
Tabla 5.2. Atípicos detectados en los meses del año 2008 (Se muestra solo el código de la serie). 97	
Tabla 5.3. Resultados totales para el ajuste del modelo del primer corte.....	103
Tabla 5.4. Predicciones sobre diciembre de 2002 comparadas con los límites superior e inferior de confianza.	106
Tabla 5.5. Resultados globales para el ajuste del modelo del segundo corte.	107
Tabla 5.6. Resultados totales para el tercer y último corte de la serie de asociados a la seguridad social en el sector de la construcción.	111
Tabla 5.7. Predicciones sobre septiembre de 2017 y datos reales, comparadas con los límites superior e inferior de confianza.....	115

ÍNDICE DE ECUACIONES.

(3.1)27

(3.2)27

(3.3)27

(3.4)27

(3.5)28

(3.6)28

(3.7)28

(3.8)30

(3.9)30

(3.10)30

(3.11)31

(3.12)31

(3.13)32

(3.14)33

(3.15)33

(3.16)33

(3.17)33

(3.18)34

(3.19)34

(3.20)34

(3.21)34

(3.22)34

(3.23)34

(3.24)35

(3.25)35

(3.26)35

(3.27)35

(3.28)	35
(3.29)	37
(3.30)	37
(3.31)	37
(3.32)	38
(3.33)	38
(3.34)	38
(3.35)	38
(3.36)	39
(3.37)	39
(4.1)	45
(4.2)	50
(4.3)	59
(4.4)	62
(4.5)	65
(4.6)	68
(4.7)	71
(4.8)	74
(4.9)	77
(4.10)	80
(4.11)	84
(4.12)	88
(4.13)	91
(5.1)	103
(5.2)	107
(5.3)	109

1 INTRODUCCIÓN Y OBJETIVOS.

El objetivo principal de este trabajo consiste en analizar series temporales pertenecientes a varios sectores económicos relacionados entre sí. Como resultado de este análisis se espera obtener información que explique el comportamiento de estos sectores, de forma global y conjunta, especialmente en torno a los años de la crisis económica de 2007-2010. Por último, se realizarán predicciones con los modelos obtenidos sobre el comportamiento futuro de alguna de las series.

A lo largo de este trabajo se van a analizar un total de 13 series temporales. Las series estudiadas pueden agruparse en 4 grandes grupos: sector de la construcción, sector inmobiliario, industria metalúrgica e indicadores económicos generales. Se elige estudiar estos tres sectores debido a la interacción entre ellos. Al estar relacionados entre sí, se espera observar como la crisis económica se propaga a través de ellos.

Para el análisis de estas series, se van a utilizar los modelos ARIMA, con la metodología Box-Jenkins y la detección automática de valores atípicos o “outliers”. Las estimaciones para estos modelos han sido hechas con dos programas: R, que es un software de código abierto especializado en estadística, y SPSS, que requiere licencia y es uno de los programas estadísticos más extendidos en el mundo empresarial. Todos los modelos han sido estimados en ambos programas, aunque los resultados presentados corresponden a los resultados obtenidos en R. No obstante, para unas pocas series y en determinados momentos se muestran resultados en SPSS para exponer el proceso en este otro programa, dada su difusión en el mundo empresarial. Por otro lado, las estimaciones finales de las predicciones han sido realizadas en SPSS.

Este primer capítulo sirve a modo de presentación del resto del trabajo. En el segundo capítulo, se expondrá brevemente la evolución de la estructura económica española a nivel general. Además, se expondrá también la situación de los sectores a los que pertenecen las series, incluyendo aportaciones al PIB y otros indicadores.

El tercer capítulo presenta el marco teórico del análisis de series temporales. Se explicará de forma breve los conceptos básicos de series temporales, y los diferentes modelos para series temporales, hasta llegar a los modelos integrados ARIMA estacionales. Después se explicará en más detalle los conceptos de análisis de intervención y de valores atípicos, presentando cada posible tipo de valor atípico en una serie temporal, acompañado de ejemplos gráficos.

El cuarto capítulo consiste en el análisis de las series temporales. El análisis de las primeras series se presentará con mayor detalle, para mostrar la aplicación directa de la teoría y el uso de los dos programas utilizados. El resto de series serán analizadas de forma sistemática. Para cada serie, se aportará como resultado del análisis un modelo ARIMA estimado, acompañado de una tabla con toda la información de los outliers detectados.

En el quinto capítulo se muestra el análisis de los resultados obtenidos tras analizar todas las series. Este capítulo está separado en dos secciones. En primer lugar, se hará un recopilatorio de todos los outliers obtenidos. Esta recopilación mostrará los efectos de la crisis, que se traducirán en una inusual concentración de atípicos en torno a los años de dicha crisis. Estudiando la distribución de los atípicos, se podrá observar la propagación de la crisis a través de los sectores estudiados. Además, se buscarán acontecimientos importantes concretos que expliquen la aparición de varios de los atípicos. Por otro lado, se utilizará una de las series analizadas para mostrar la potencia de los modelos ARIMA a la hora de generar predicciones. Esta serie será “cortada” en varios puntos del tiempo, en los cuales se haya detectado algún valor atípico. Sobre estos cortes se generarán predicciones basadas en modelos ARIMA, que se compararán con los datos reales. Por último, se generarán predicciones para la serie total analizada, y se compararán con los datos de dicha serie publicados posteriormente a la realización de este trabajo. A continuación, se dedicará el último capítulo a hablar sobre las conclusiones que se pueden extraer de este trabajo.

En último lugar, se incluye un anexo mostrando el código de R utilizado para el análisis de las series.

2 ESTRUCTURA ECONÓMICA ESPAÑOLA. EVOLUCIÓN DEL SECTOR DE LA CONSTRUCCIÓN, SECTOR INMOBILIARIO Y DE LA INDUSTRIA METALÚRGICA.

2.1 Estructura de la economía española.

La economía española está constituida por cuatro sectores, los cuales son: agricultura y pesca, industria y energía, construcción y sector servicios. A continuación, se expondrá brevemente la evolución y situación actual de estos sectores. La figura 2.1 muestra un gráfico comparativo donde se puede apreciar el peso de cada sector en el PIB español, y su evolución durante los últimos veinte años.

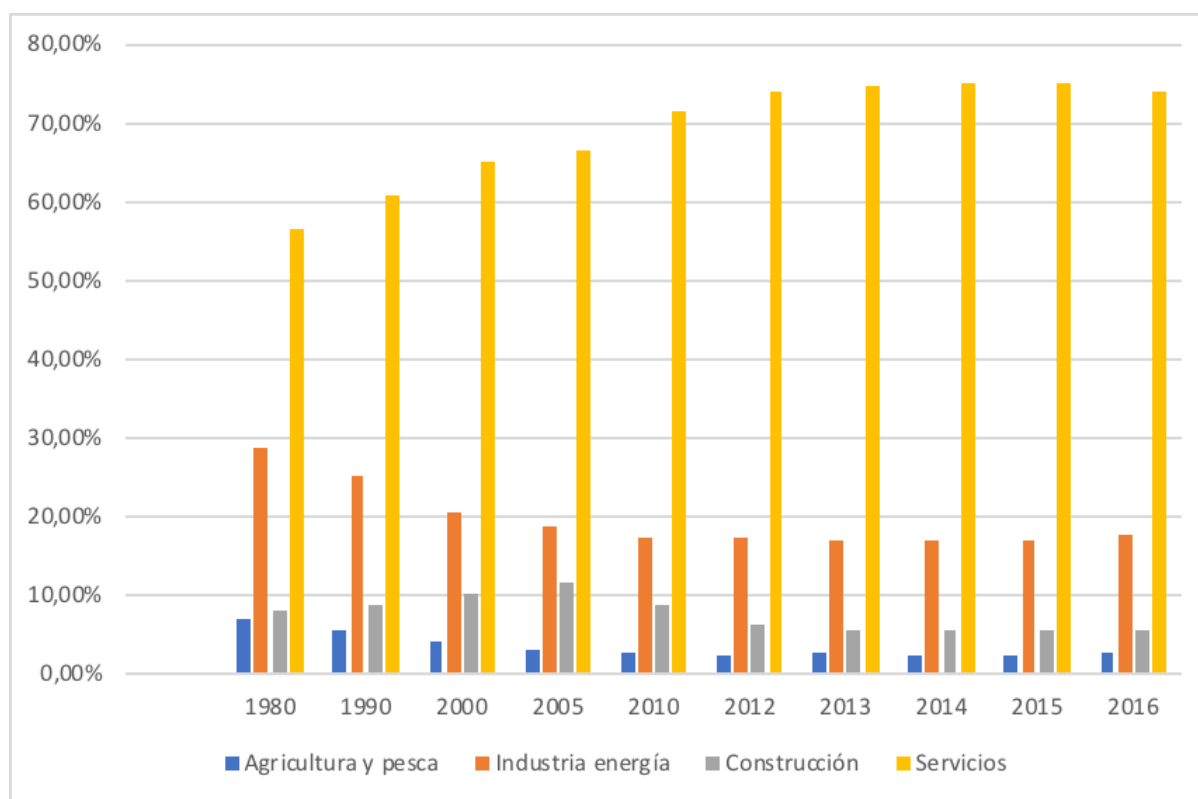


Figura 2.1. Evolución del peso porcentual de los distintos sectores en España sobre el PIB desde el año 1980. Fuente: Elaboración propia a partir de datos extraídos del INE.

Observando la evolución desde 1980 hasta 2016, se puede apreciar un claro incremento del sector terciario o servicios, que coincide con el proceso de “desagrarización” que sufrió España a lo largo de estos años. Por otra parte, el sector industrial descende su contribución al PIB de forma gradual hasta los años de la crisis, momento en el cual se mantiene estables

a niveles mínimos, sin mostrar signos importantes de recuperación. En 2016, observamos que el sector de la industria ocupa el segundo lugar en importancia con un 17,8%, seguido por la construcción con un 5,6%. A raíz de estos indicadores, se puede vislumbrar la reestructuración que ha atravesado la economía española en las últimas tres décadas. el sector de la industria ha sufrido un desplome del 3.5% en su contribución al PIB desde el comienzo del nuevo milenio.

Para entender esta reestructuración de la economía, es necesario buscar causas subyacentes para el descenso de la contribución al PIB del sector industrial. En la figura 2.2 se puede observar la contribución porcentual al empleo por parte de los cuatro sectores principales.

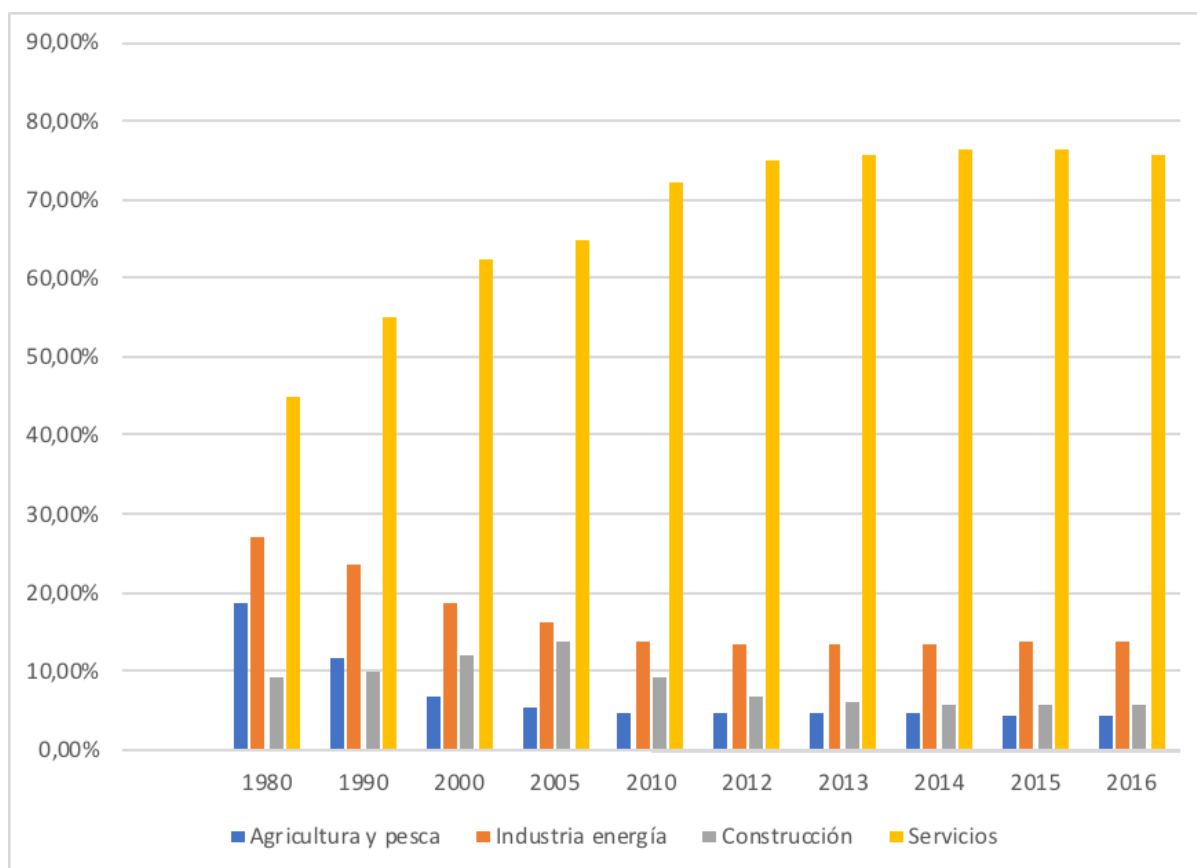


Figura 2.2. Evolución de la estructura porcentual del empleo en España desde el año 1980 (en porcentaje). Fuente: Elaboración propia a partir de datos extraídos del INE.

Hasta el año 2005, se muestra como el sector de la industria proporciona un alto porcentaje relativo del empleo total, situándose como el segundo sector que mayor cantidad de empleo ofrece. Sin embargo, a partir de este año, y en adelante a la crisis económica, se ve un

decrecimiento en la aportación al porcentaje de empleo. Aun así, se mantiene como el segundo sector que mayor ocupación proporciona.

Es importante prestar atención brevemente a la inversión en I+D+i que se realiza en este sector. En el año 2010 se hizo una inversión del 0,84% del valor de la producción, suponiendo menos de la mitad de media del gasto que realiza la UE en otros países como en Alemania. La figura 2.3 muestra el porcentaje de gasto de inversión desgranado por cada sector de la industria, acompañado de la diferencia con respecto a la media del gasto europeo en estos sectores. Podemos apreciar que, en global, el gasto es efectivamente menor que la media europea. Esta falta de inversión en el sector está directamente relacionada con el descenso en el empleo ofrecido por este sector.

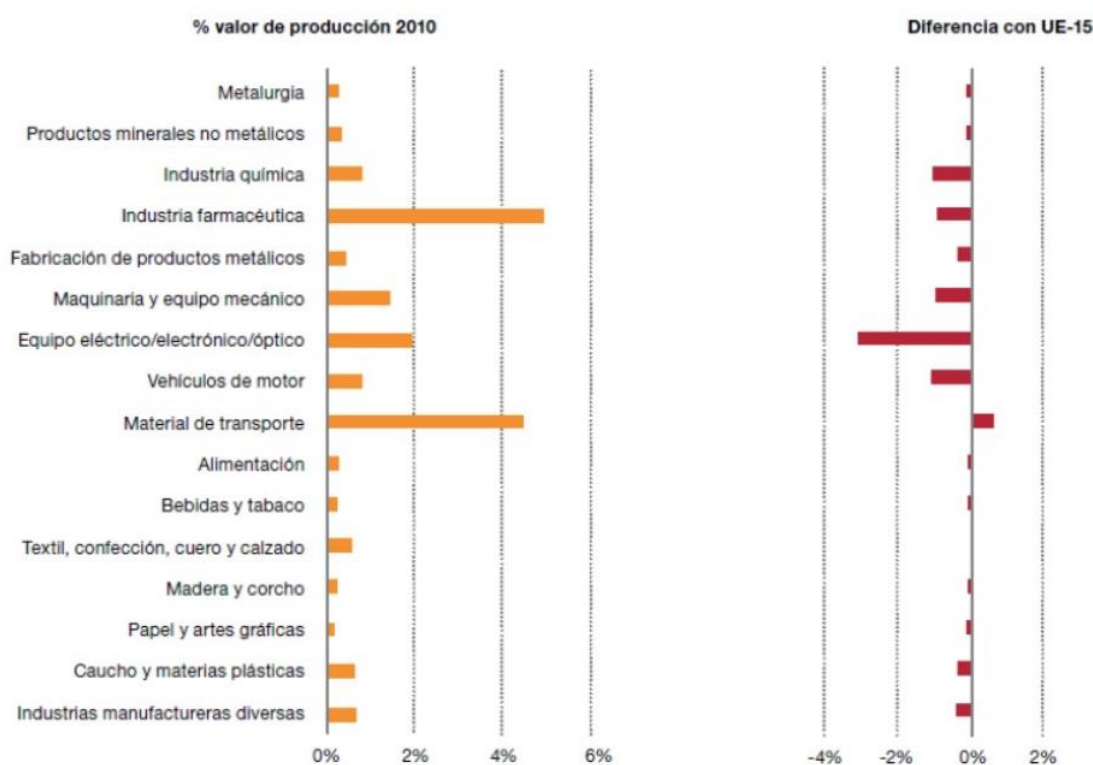


Figura 2.3. Inversión en I+D en la industria española en comparación con UE15 en el año 2010 (en porcentaje). Fuente: Eurostat.

Atendiendo a la variación anual del IPI (Índice de Producción Industrial) entre 1983 y 2018, visible en la figura 2.4, se muestra un comportamiento relativamente estable en el sector hasta los años de la crisis. Durante la crisis se observa un fuerte descenso en la actividad del sector. Esta caída no comienza a mostrar signos de recuperación hasta casi tres años más tarde. A partir de dicho punto, la volatilidad del sector aumenta notablemente.

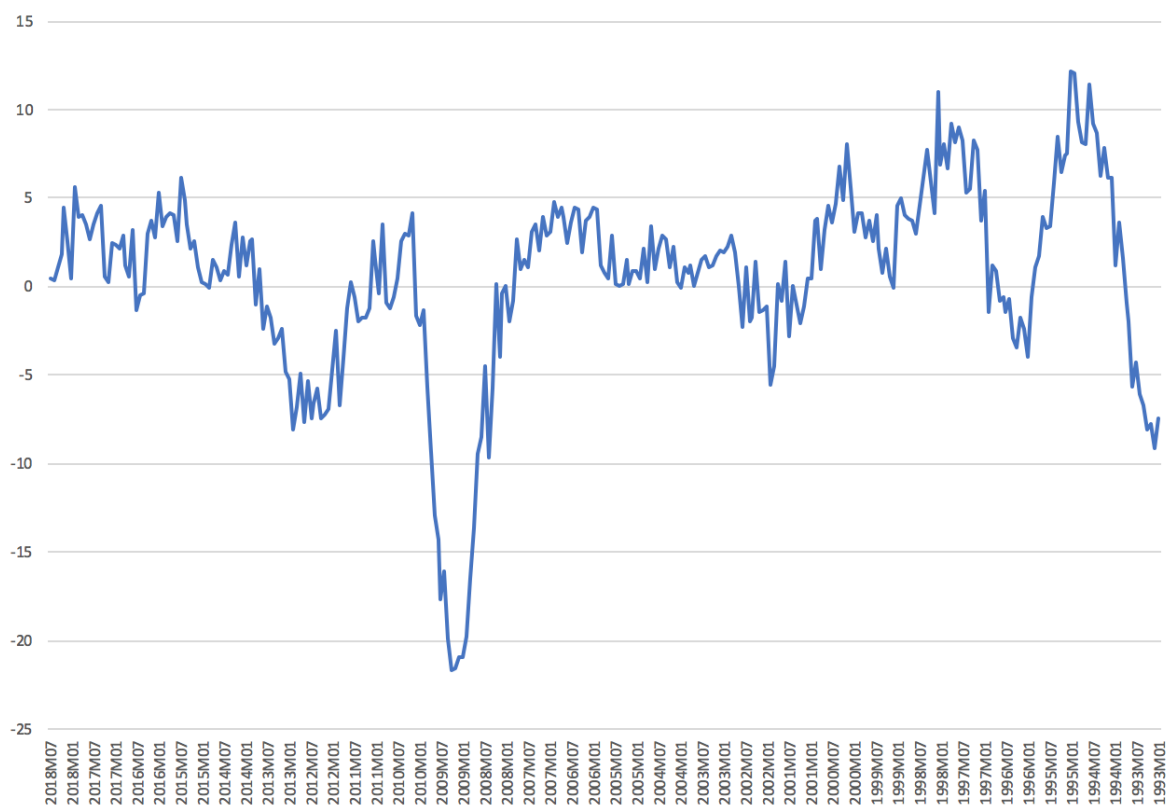


Figura 2.4. Crecimiento del Índice de Producción Industrial, desde 1993 hasta 2018 (en porcentaje). Fuente: Elaboración propia a partir de datos extraídos del INE.

A continuación, se procede a estudiar los sectores implicados en este trabajo: construcción, inmobiliario y metalurgia.

2.2 Evolución histórica del sector de la construcción.

El sector de la construcción en España tiene una gran importancia, debido a su papel determinante en la economía del país, así como por la gran cantidad de empleo que genera. Las aportaciones del sector de la construcción al PIB y al empleo son, en ambos casos, de entorno al 10%. Según la Organización Internacional del Trabajo, en el sector de la construcción predomina el trabajo realizado por y a través de pequeñas y medianas empresas. Además, la actividad en este sector suele tener un carácter mayoritariamente local, en contraposición a la globalización que afecta a todos los sectores económicos hoy en día.

Debido a su importante papel en la economía española, cuando se desencadenó la última crisis económica en 2007, varios indicadores relacionados con la construcción se vieron especialmente afectados. En el caso de los empleados del sector de la construcción (figura

2.5) se puede observar un incremento constante hasta el 2007, año en el que se aprecia un fuerte descenso, coincidiendo con la crisis económica. En 2013 comienza a estabilizarse, mostrando leves signos de recuperación en los últimos años.

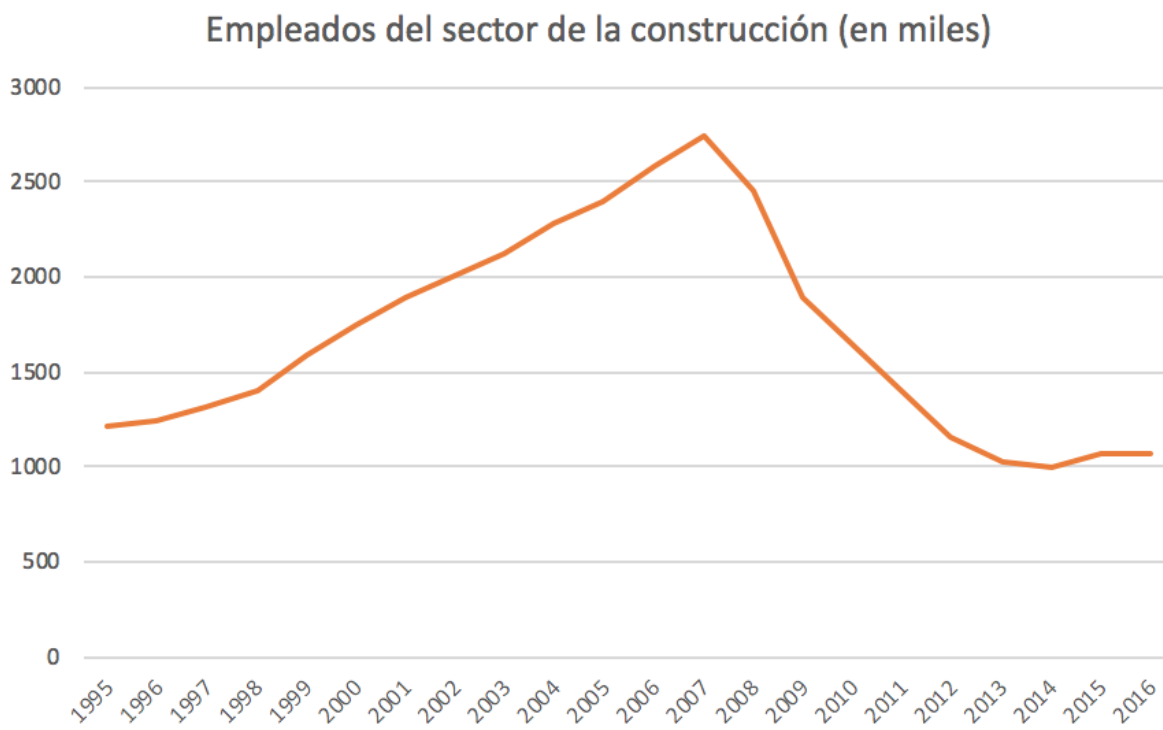


Figura 2.5. Evolución del número de empleados del sector de la construcción desde 1995 hasta 2016 (en miles). Fuente: Elaboración propia a partir de datos extraídos del INE.

Una tendencia similar puede ser observada en el consumo de cemento en España (figura 2.6), relacionado directamente con el sector de la construcción por su uso en obras. El descenso en el uso cemento fue especialmente drástico, debido a la paralización de una gran cantidad de obras. Además, tal y como se verá en el análisis de las series, los índices de precios en el sector de la construcción continuaron aumentando a pesar de la llegada de la crisis.

Haciendo uso de la figura 2.1, se observa como este sector alcanza su punto de mayor contribución al PIB en los años de la crisis, momento a partir del cual se desploma. Este desplome acarrea un descenso de la contribución en torno al 5%, lo que supone casi la mitad de la contribución que realizaba durante el auge la burbuja inmobiliaria previa a la crisis. Por otra parte, observando de nuevo el gráfico 2.2, vemos una tendencia similar en el porcentaje de empleo aportado por este sector. No obstante, la caída en este aspecto es aún

más acusada bajando de entorno al 15% a valores cercanos al 5%. Este desplome en la aportación de empleo por parte de este sector es un claro indicador de como la burbuja inmobiliaria se trasladó al sector de la construcción, contaminándolo, y haciendo que los efectos de la crisis se recrudecieran en el sector de la construcción.

Como conclusión, se puede decir que el sector de la construcción fue, junto al sector industrial, uno de los más afectados por la crisis, mostrándose descensos en la actividad y productividad muy significativos.

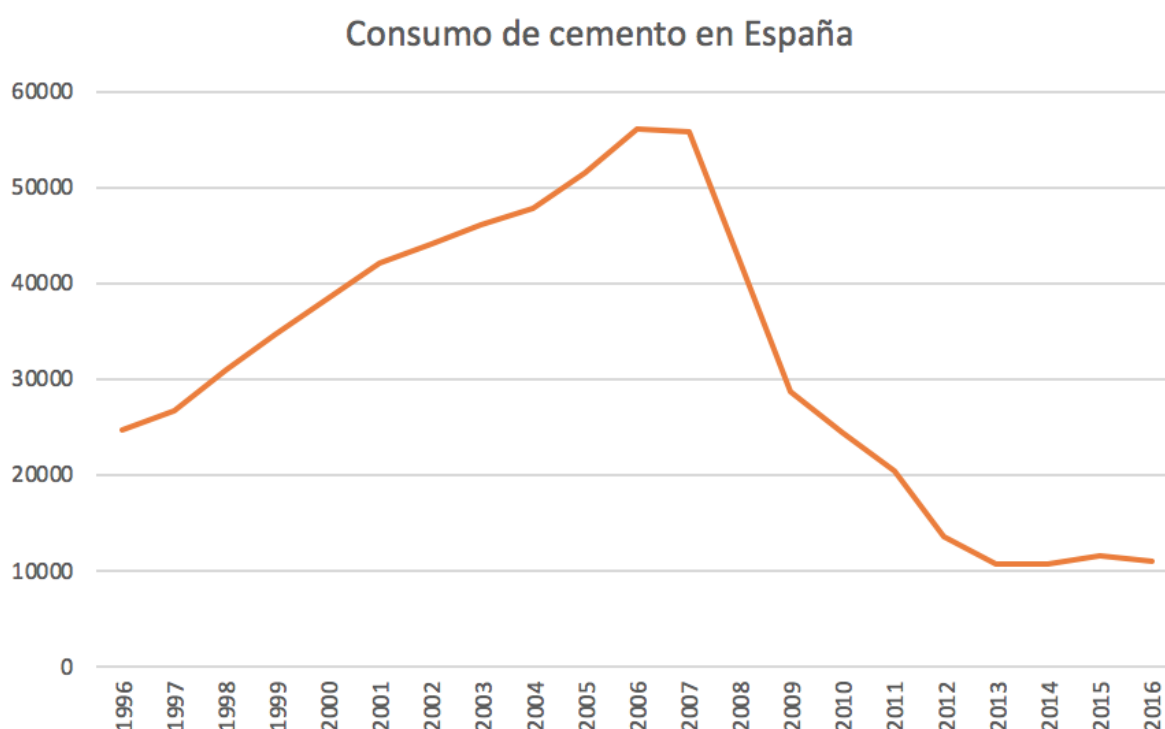


Figura 2.6. Consumo de cemento en España desde 1996 hasta 2016 (en miles de toneladas). Fuente: Elaboración propia a partir de datos extraídos del INE.

2.3 Evolución del sector inmobiliario.

Directamente relacionado con el sector de la construcción se encuentra el sector inmobiliario. Aunque el sector de la construcción abarca una gran cantidad de proyectos, la construcción de viviendas o inmuebles habitables es una de sus principales fuentes de actividad. Es aquí donde se relaciona con el sector inmobiliario, encargado de distribuir el producto “vivienda” a los consumidores finales. La burbuja inmobiliaria, cuyas causas se expondrán a continuación, fue el principal agravante de la crisis en nuestro país, y su explosión produjo graves consecuencias en el ya expuesto sector de la construcción. Una burbuja inmobiliaria puede definirse como una especulación existente sobre bienes inmuebles en una economía. En el caso de España, las dimensiones que alcanzó dicha

burbuja fueron desmesuradas, y contaminó a los sectores adyacentes al inmobiliario, así como a los bancos involucrados en la actividad del sector.

Existen varios indicadores económicos que muestran el auge y desplome de esta burbuja. Por un lado, en la figura 2.7, se puede observar la evolución del precio de la vivienda en España. Es claramente apreciable como el precio de la vivienda se incrementa constantemente desde el año 2000, alcanzando su máximo en 2007. Una situación insostenible para la sociedad y economía española del momento, que colapso bajo los efectos de la crisis económica originada en Estados Unidos

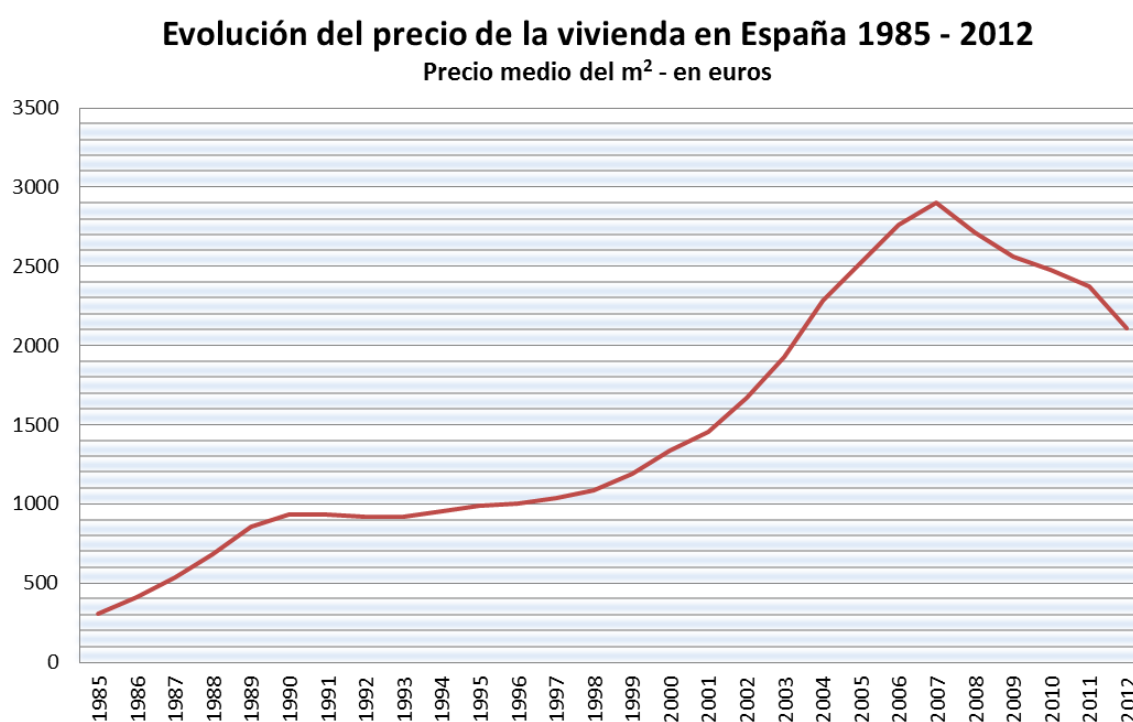


Figura 2.7. Evolución del precio de la vivienda en España desde 1985 hasta 2012 (precio medio del m² en euros). Fuente: INE.

Además, la oferta y demanda de la vivienda en España también se vio severamente afectada. En la figura 2.8 se muestra cómo la oferta de viviendas nuevas es superior a la demanda existente. Este hecho desemboca en que a partir de 2007 se detenga el inicio de nuevas obras. Además, el número de obras terminadas se mantiene inalterado durante varios años al no haber demanda. A partir del 2015 se empieza a observar una sensible recuperación en las ventas de viviendas nuevas.

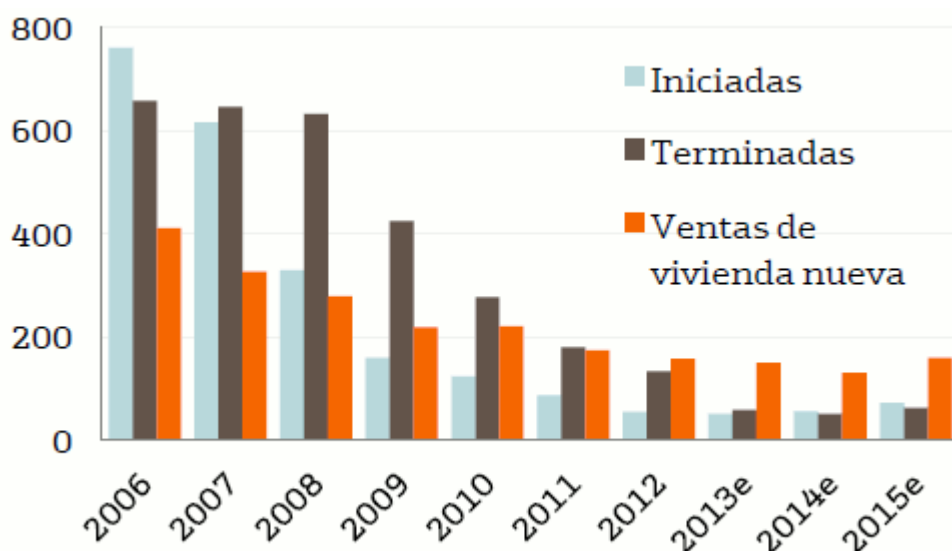


Figura 2.8. Oferta y demanda de vivienda en España desde el 2006 hasta el 2015. Fuente: INE y Banco de España realizada por Bankinter.

Algunos de los grandes afectados por la burbuja inmobiliaria fueron las cajas de ahorro, las cuales mostraron un comportamiento temerario, al conceder créditos de manera indiscriminada. Esta situación desemboca en que muchas de estas cajas quebraron. A raíz de estas quiebras, se llevó a cabo una reconstrucción bancaria de mano de la UE, que prestó grandes sumas de dinero para llevar a cabo el rescate de estas cajas y bancos. Aun así, estos préstamos no fueron suficientes para solucionar la situación, y el dinero correspondiente aún no ha sido devuelto por parte de estas entidades. Uno de los ejemplos más mediáticos es el caso Bankia.

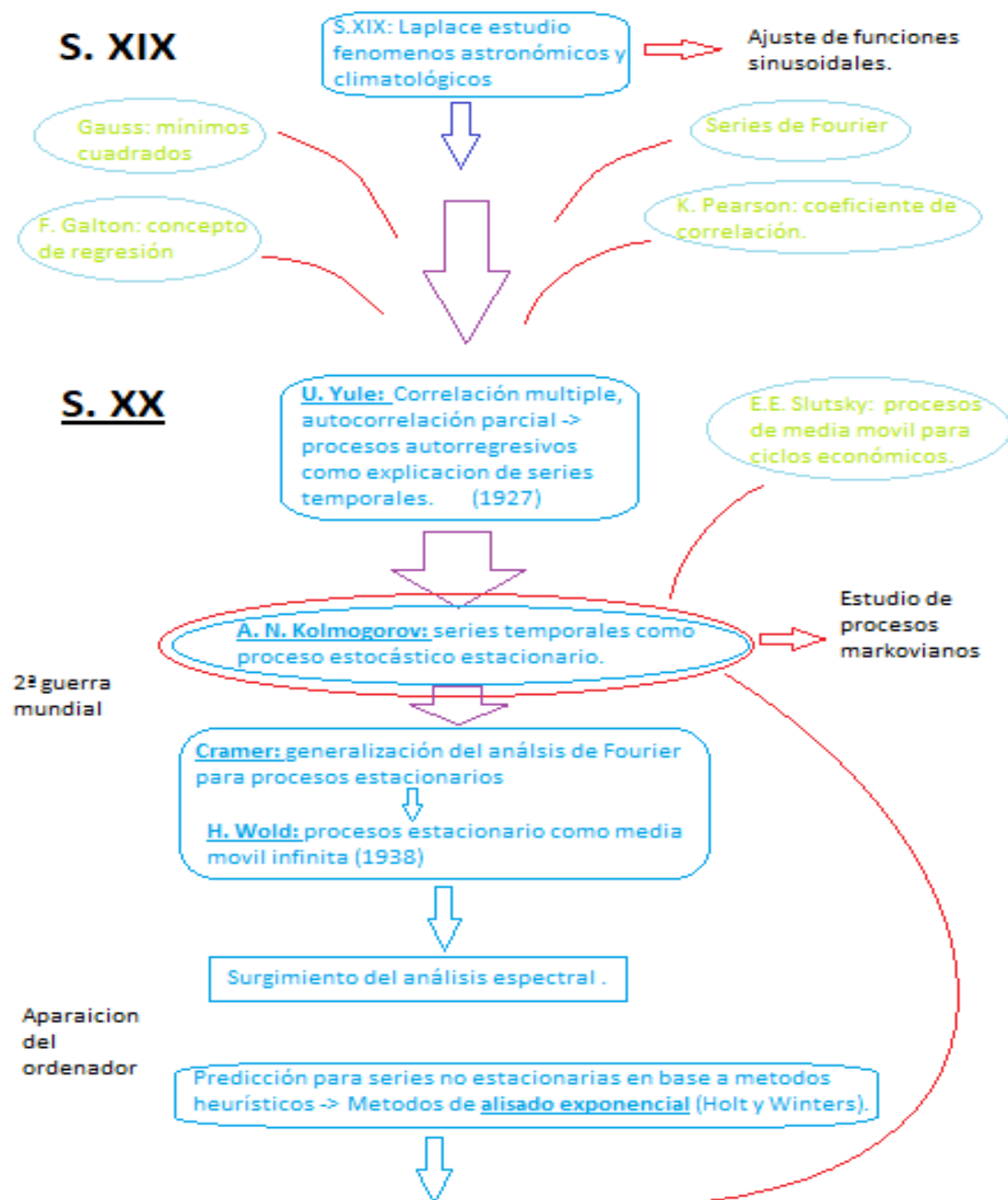
2.4 Industria metalúrgica.

La industria metalúrgica ha sido históricamente de gran importancia en nuestro país, y una de las más destacadas en el sector industrial. En concreto, el acero y las aleaciones férreas suponen una de las principales fuentes de materias primas para el sector de la construcción. Como ya se ha expuesto en el apartado de la estructura económica española, el sector de la industria supone un componente esencial en la economía de nuestro país. La inclusión del estudio de este sector en el trabajo tiene como objetivo determinar si la burbuja inmobiliaria que afectó al sector de la construcción llegó a propagarse hasta la industria metalúrgica, así como para comparar los efectos de la crisis con el resto de sectores estudiados. El comportamiento de esta industria sigue los patrones ya observados para el sector industrial

3 INTRODUCCIÓN TEÓRICA AL ANÁLISIS DE SERIES TEMPORALES.

3.1 Evolución histórica del análisis de series temporales.

Con el fin de situar en el tiempo los orígenes y proliferación de las técnicas de análisis de series temporales que se van a usar en este trabajo, se muestra a continuación un gráfico de realización propia que ilustra brevemente la evolución histórica de esta disciplina.



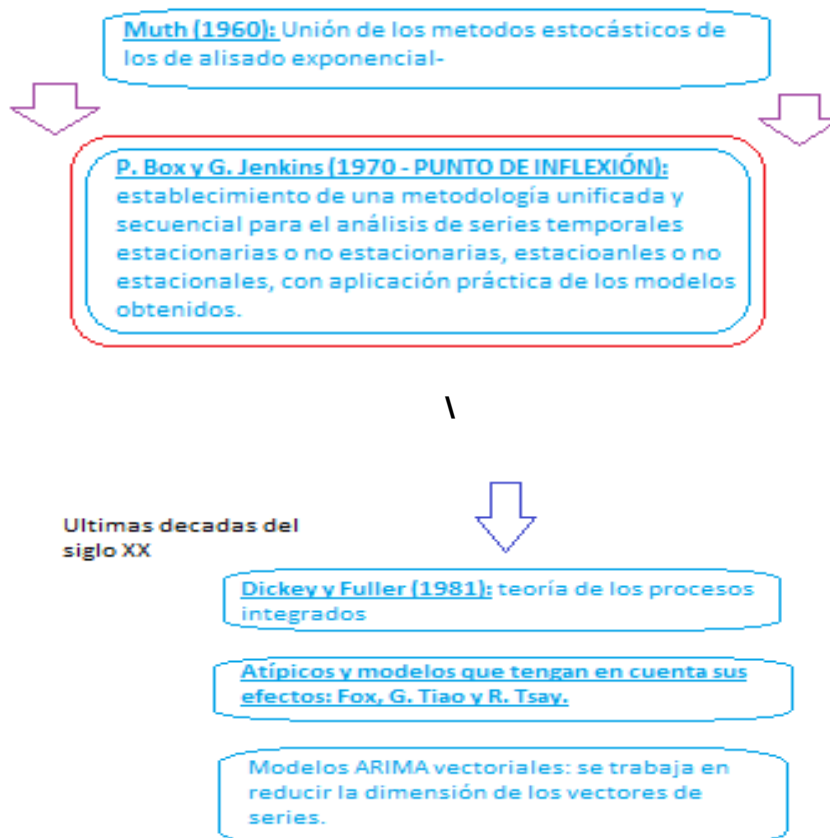


Figura 3.1. Evolución histórica del análisis de series temporales.

3.2 Concepto de serie temporal.

Se define como serie temporal un conjunto de datos indexados en orden cronológico. Estos datos son generalmente medidos de forma periódica. Las series temporales se representan mediante la notación z_t , donde z es la serie temporal y t el periodo de medición.

$$z_t$$

La serie temporal por lo tanto será el conjunto de datos z_1, z_2, z_3, \dots

3.3 Características de una serie temporal. Definición de serie estacionaria.

En este apartado se pasa a describir las tres características principales de una serie temporal que deben ser tenidas en cuenta para su estudio y análisis.

- **Tendencia:** una serie temporal presentará tendencia (local o en toda su extensión), si sus valores aumentan o disminuyen con el tiempo. En ausencia de tendencia, la serie temporal gravitará en torno a un determinado valor
- **Volatilidad:** la volatilidad de una serie temporal hace referencia a la disparidad o similitud de los valores de sus observaciones. Un nivel de volatilidad estable dará lugar a una serie temporal homocedástica. En caso contrario, la serie será heterocedástica.
- **Ciclo estacional** > La estacionalidad es un fenómeno que tiene lugar cuando las series temporales se observan más de una vez al año. Por ejemplo, una serie con datos mensuales es normal que presente comportamientos similares en cada ciclo anual.

Se define como **serie estacionaria** aquella serie temporal que no presenta tendencia ni ciclo, y es homocedástica. Si una serie no es estacionaria, se puede llevar a cabo una serie de transformaciones que serán descritas en este capítulo hasta transformarla en estacionaria. Este proceso es necesario para poder aplicar los modelos ARIMA (que se definirán más tarde)..

3.4 Estructuras de dependencia: función de autocorrelación simple (FAS) y función de autocorrelación parcial (FAP).

La idea detrás de conocer las funciones de autocorrelación simple y parcial de una serie temporal es que estas nos permiten determinar las relaciones de dependencia lineal entre las observaciones que la componen. Estas relaciones hacen referencia a cómo afectan las observaciones pasadas sobre las futuras. La información presentada por ambas funciones de autocorrelación es exclusiva y complementaria, de tal forma que conociendo ambas funciones para una serie temporal, es posible identificar el tipo de modelo que explicaría su comportamiento.

- **Función de autocorrelación simple (FAS):** proporciona los coeficientes de autocorrelación lineal entre dos observaciones de la serie temporal separadas por un número concreto de periodos. En otras palabras, indica cómo influye el valor de una observación de forma lineal sobre el valor del resto de observaciones futuras. No obstante, una observación separada “k” retardos de otra, influenciara sobre esta última a través de las observaciones comprendidas en esos “k” retardos. La FAS

estará formada por la sucesión de estos coeficientes de autocorrelación, cuyo valor debe estar comprendido entre -1 y 1. El primer coeficiente indica la influencia de una observación sobre la del instante inmediatamente posterior, el segundo sobre la de dos instantes después, el tercero sobre la de tres... y así sucesivamente hasta acabar con todo el rango de observaciones obtenidas. Una vez obtenida la función, puede ser representada gráficamente mediante gráficas de “palos”.

- **Función de autocorrelación parcial (FAP):** La función de autocorrelación parcial proporciona los coeficientes de correlación lineal entre observaciones de forma directa, es decir como influencia una variable sobre otra separada “k” retardos, sin tener en cuenta la influencia a través de las variables intermedias entre esos “k” retardos. Sigue las mismas reglas matemáticas y de representación que la FAS.

A continuación, se muestra una imagen ilustrando los posibles tipos de influencia de una observación previa sobre el resto de las observaciones. De esta forma, vemos como por ejemplo la primera observación puede afectar a la tercera a través de la segunda, o bien hacerlo directamente. Por lo tanto, los coeficientes de autocorrelación tanto simple como parcial entre la primera y tercera observación serán no nulos.

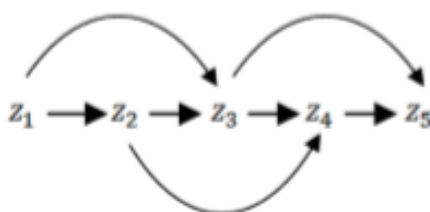


Figura 3.2. Posibles formas de influencia de las observaciones sobre las demás.

En las gráficas de las funciones de autocorrelación simple y parcial, se representan dos bandas de significación. Si un palo de la gráfica supera alguna de estas bandas se considera que el palo es significativo, y que por lo tanto la correlación que indica entre dos observaciones también lo es. La siguiente figura muestra el ejemplo de las gráficas de la FAS y la FAP de una serie temporal.

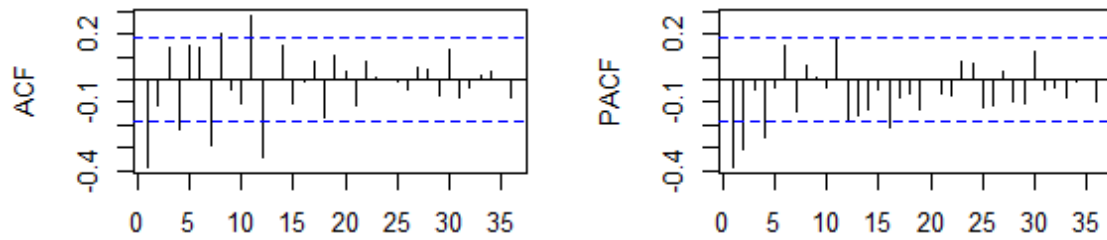


Figura 3.3. Gráficas FAS y FAP de una serie temporal.

3.5 Ruido Blanco.

En términos de series temporales, un ruido blanco puede ser considerado como una serie temporal en la que no hay ningún tipo de correlación entre ninguna de las variables que conforman la serie. El gráfico de secuencia de un ruido blanco no presenta ningún tipo de estructura, y tanto como su FAS y su FAP no poseen palos significativos. A la hora de identificar un ruido blanco, es necesario observar la FAS y la FAP, ya que es imposible de determinar con claridad simplemente a partir del gráfico de secuencia.

El concepto de ruido blanco es importante en el análisis de series temporales. Como se explicará en los siguientes apartados, una de las formas de comprobar si un modelo matemático explica correctamente el comportamiento de una serie temporal es observar los residuos que genera al ajustar dicho modelo. Si estos residuos se comportan como un ruido blanco, será un claro indicador de que el modelo es bueno. La siguiente imagen muestra las gráficas de la FAS y la FAP de un ruido blanco.

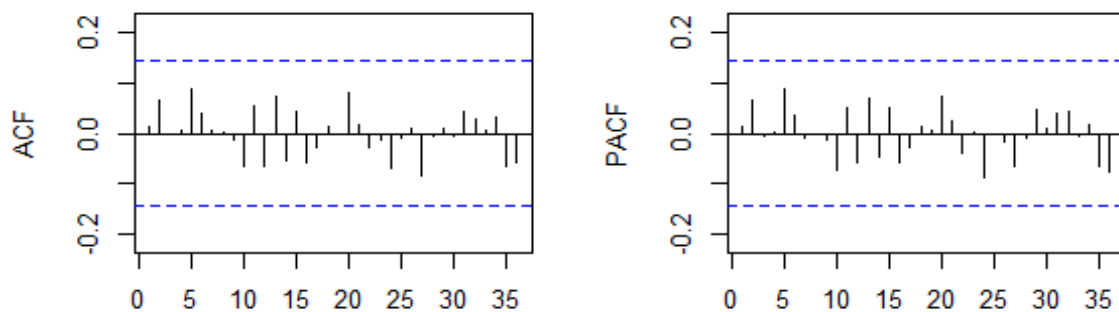


Figura 3.4. Gráficas FAS y FAP de un ruido blanco.

3.6 Operadores matemáticos utilizados en el análisis de series temporales.

En la notación matemática utilizada en el análisis de series temporales se utilizan una serie de operadores matemáticos que se describen a continuación:

- **Operador de retardo:** B es el operador de retardo, que queda definido como “un operador lineal que aplicado a una función temporal, proporciona esta misma función retardada un periodo” (Peña 2010)

Aplicando este operador a una observación de la serie, se obtiene el siguiente resultado:

$$Bz_t = z_{t-1} \quad (3.1)$$

$$B^k z_t = z_{t-k} \quad (3.2)$$

Este operador resulta de gran utilidad a la hora de expresar las ecuaciones de los modelos ARIMA para series temporales que se describirán más adelante.

- **Operador de diferencia:** ∇ es el operador de diferencia, que queda definido de la siguiente forma:

$$\nabla = z_t - z_{t-1} = (1 - B)z_t \quad (3.3)$$

- **Operador de diferencia estacional:** ∇_s es el operador de diferencia estacional, definido de la siguiente forma:

$$\nabla_s = z_t - z_{t-s} = (1 - B^s)z_t \quad (3.4)$$

Al aplicarlo a la serie temporal obtenemos el mismo resultado que con el operador de diferencia regular, salvo que la diferencia es entre las observaciones separadas k retardos. Este operador se utiliza para eliminar los ciclos estacionales presentes en la serie, de estacionalidad s .

3.7 Modelos autorregresivos, AR(p).

Los modelos autorregresivos explican las series temporales en base a relaciones de autorregresión lineales entre las observaciones. Según estos modelos, cada observación se construye como la combinación lineal de las observaciones anteriores multiplicadas por determinados factores, más una perturbación aleatoria llamada innovación, que sigue las reglas de un ruido blanco. En los modelos AR, poseen un determinado orden p . El orden del modelo AR determina la cantidad de observaciones (inmediatamente anteriores) con las que se construye una determinada. De esta forma, un modelo de orden uno, AR (1), se expresa de la siguiente forma:

$$z_t = \phi_1 z_{t-1} + a_t \quad (3.5)$$

donde:

$$\phi_1 \in (-1,1)$$

para que la serie temporal sea estacionaria.

Un modelo AR (p) general sigue la siguiente ecuación:

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + a_t \quad (3.6)$$

Utilizando el operador de retardo, se puede simplificar esta expresión para obtener la ecuación que se usa normalmente para expresar este tipo de modelos:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) z_t = a_t \quad (3.7)$$

Para que la serie sea estacionaria, todas las raíces del polinomio en B deben estar fuera del círculo unidad.

Este tipo de modelos se caracterizan por propagar durante largos periodos de tiempo los impactos externos sobre la serie (absorción lenta del impacto). A continuación, se muestra las gráficas de la FAS y la FAP de un modelo AR (1), en función del signo del coeficiente ϕ_1 . También se muestran dichas gráficas para una de las posibles opciones de modelo AR (2). (Peña 2010)

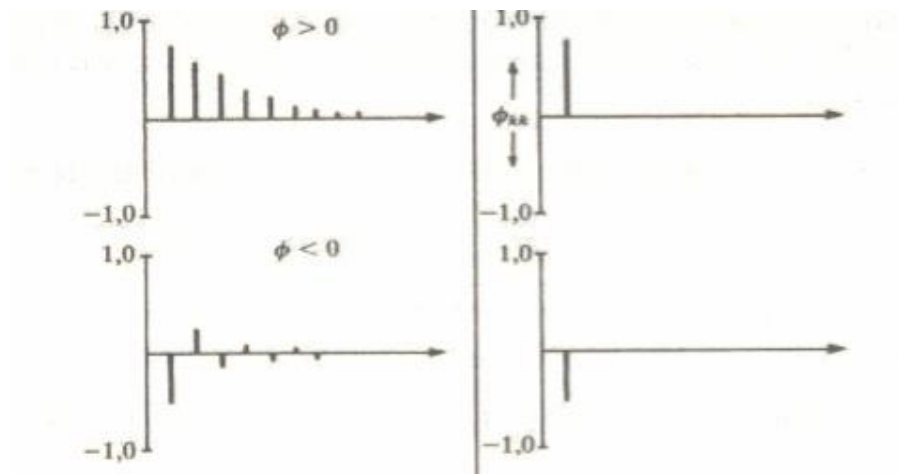


Figura 3.5. FAS y FAP para modelos AR (1). (Peña 2010).

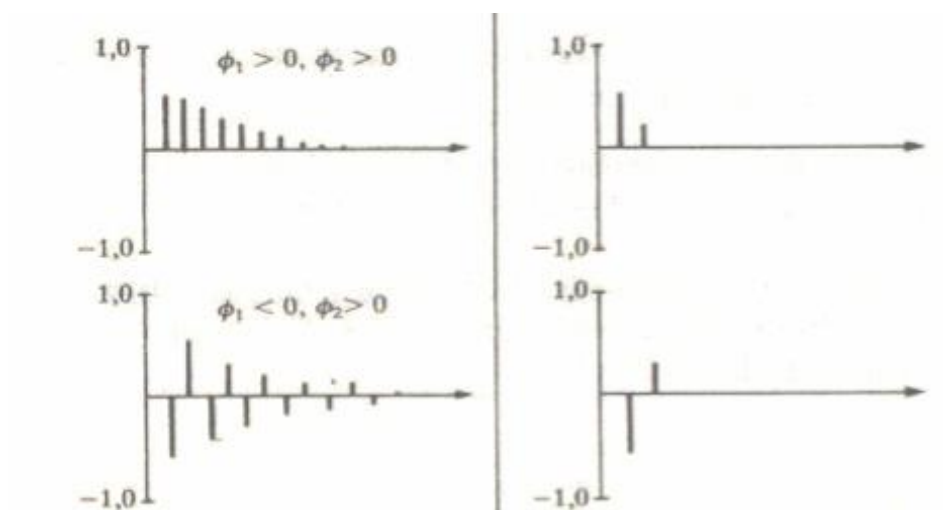


Figura 3.6. FAS y FAP para algunos modelos AR (2). (Peña 2010).

3.8 Modelos de media móvil, MA (q).

Los modelos de media móvil, MA, por sus siglas en inglés, “moving average”, establecen que cada observación se construye como una combinación lineal de la innovación propia de dicha observación, sumada a las innovaciones de un determinado número de observaciones inmediatamente anteriores, multiplicados por sus específicos coeficientes. Los modelos MA presentan un determinado orden q , que determina la cantidad de innovaciones de

observaciones inmediatamente anteriores que afectan a una observación. De esta forma, un modelo MA de orden uno, MA (1), presenta la siguiente ecuación:

$$z_t = a_t - \theta_1 a_{t-1} \quad (3.8)$$

$$\theta_1 \in (-1,1)$$

En consecuencia, un modelo MA (q) general se expresa de la siguiente forma:

$$z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (3.9)$$

$$z_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t \quad (3.10)$$

Para que la serie sea invertible (es decir, que se pueda escribir como un proceso AR infinito), todas las raíces del polinomio en B deben estar fuera círculo unidad.

Este tipo de modelos se caracterizan una rápida absorción de los impactos externos sobre la serie. A continuación, se muestra las gráficas de la FAS y la FAP de un modelo MA (1), en función del signo del coeficiente θ_1 . También se muestran una opción de dichas gráficas para una de las posibles opciones de modelo MA (2) (Peña 2015),

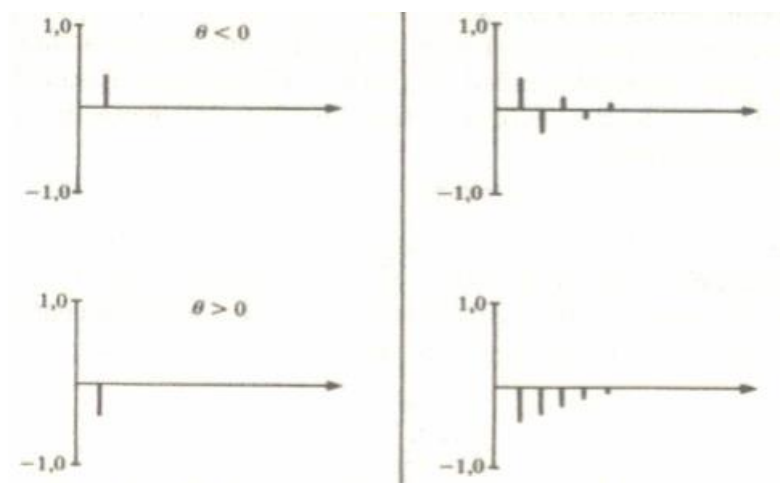


Figura 3.7. FAS y FAP para modelos MA (1). (Peña 2010).

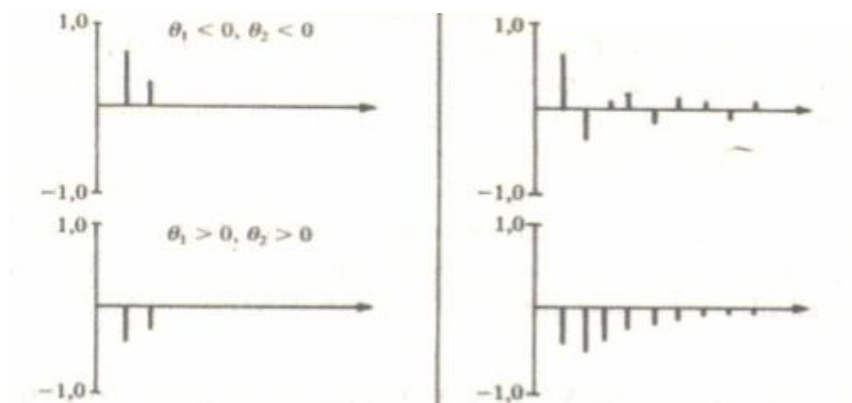


Figura 3.8. FAS y FAP para algunos modelos MA (q). (Peña 2010).

3.9 Modelos ARMA (p, q).

Con el objetivo de poder contar con las propiedades de los modelos AR y MA en un solo modelo unificado, se definen los modelos ARMA. La construcción de estos modelos se basa en unir la estructura y propiedades de sus dos componentes. La expresión general para un modelo ARMA (p, q) es la siguiente:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) z_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t \quad (3.11)$$

compactando los polinomios de ambos lados en función de B se obtiene:

$$\phi(B) z_t = \theta(B) a_t \quad (3.12)$$

Estos modelos siguen las mismas reglas que los explicados anteriormente. A continuación, se muestra la FAS y la FAP de algunos modelos ARMA de ejemplo.

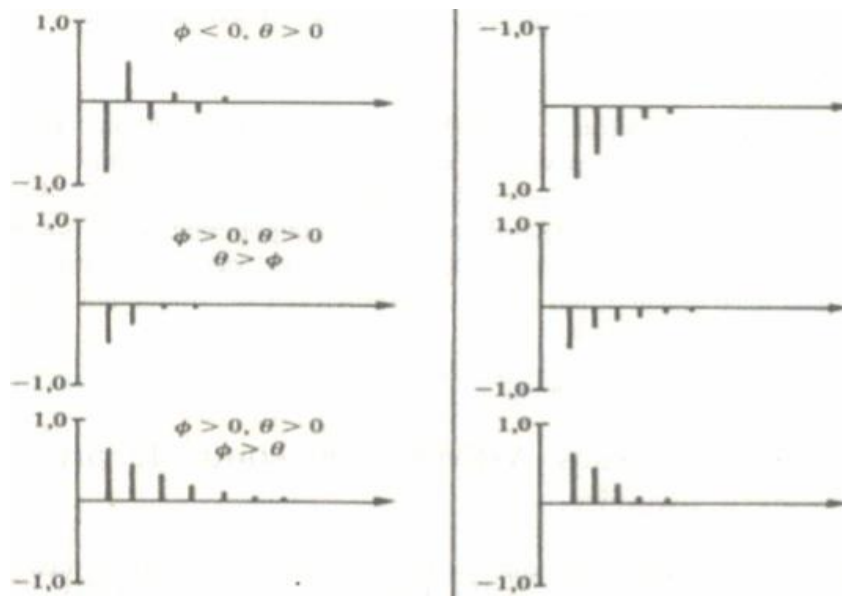


Figura 3.9. FAS y FAP de modelos ARMA. (Peña 2010).

3.10 Estacionalidad.

La estacionalidad es un fenómeno que tiene lugar cuando las series temporales se observan más de una vez al año (meses, trimestres, cuatrimestres, semestres...). Dicho fenómeno consiste en que las observaciones de un ciclo afectan de forma directa a las observaciones equivalentes de los ciclos posteriores. De esta forma, los modelos ARMA también son capaces de explicar estas relaciones estacionales entre las observaciones, dando lugar a modelos ARMA estacionales. Estos modelos estacionales siguen las mismas reglas y estructuras que los modelos ARMA regulares, salvo que las estructuras de dependencia siguen un orden de estacionalidad s . Este orden de estacionalidad indica la cantidad de observaciones que hay entre las dos observaciones estacionalmente relacionadas. Por ejemplo, $s = 12$ implica un ciclo mensual, en el que las observaciones de cada mes de un determinado año influyen sobre las de los mismos meses de los años posteriores. La expresión para los modelos ARMA estacionales es la siguiente:

$$(1 - \Phi_1 B^s - \dots - \Phi_p B^{sp})(1 - \phi_1 B - \dots - \phi_p B^p)z_t = (1 - \Theta_1 B^s - \dots - \Theta_q B^{sq})(1 - \theta_1 B - \dots - \theta_q B^q)a_t \quad (3.13)$$

3.11 Modelos integrados ARIMA (p, d, q) x (P, D, Q) s.

Estos modelos recogen todo lo explicado anteriormente para explicar series con componente tanto regular como estacional. La integración hace referencia a que también recogen las posibles diferencias, regulares y estacionales, que puede necesitar una serie para ser transformada en estacionaria. Estas diferencias están reflejadas respectivamente en los parámetros d y D de la notación del modelo. La expresión general de estos modelos se muestra a continuación:

$$(1 - \Phi_1 B^S - \dots - \Phi_P B^{SP})(1 - \phi_1 B - \dots - \phi_p B^p) \nabla_S^D \nabla^d z_t = (1 - \theta_1 B^S - \dots - \theta_Q B^{SQ})(1 - \theta_1 B - \dots - \theta_q B^q) a_t \quad (3.14)$$

En la forma compactada podemos escribir:

$$\Phi(B) \phi(B) \nabla_S^D \nabla^d z_t = \Theta(B) \theta(B) a_t \quad (3.15)$$

Si todos los términos de esta expresión se agrupan al lado derecho de la ecuación se obtiene una notación más compacta:

$$z_t = \psi(B) a_t \quad (3.16)$$

3.12 Intervenciones.

La idea principal detrás del concepto de intervención es la siguiente: en la realidad, ocurren eventos que modifican los valores de las series temporales en momentos determinados. En dichos instantes, el valor de la serie será difícilmente explicable mediante el modelo ARIMA que explica el resto de la serie. En consecuencia, será necesario incluir una variable explicativa que modele este efecto determinista sobre la serie temporal, con el objetivo de mejorar la estimación de los parámetros del modelo. Una intervención queda definida de la siguiente forma:

$$z_t = \psi(B) a_t + x_t \quad (3.17)$$

donde x_t es la intervención.

Los eventos que ocurren en el tiempo y afectan a la serie pueden ser de distinta naturaleza, siendo los más comunes acontecimientos puntuales que se diluyen rápidamente en el tiempo, o hechos que cambian el comportamiento de la serie de forma permanente una vez que

acontecen. A su vez, estas dos naturalezas dan lugar a los dos tipos de variables explicativas más comunes a la hora de generar intervenciones: variables impulso y variables escalón

Las variables impulso se utilizan para modelar acontecimientos concretos y puntuales en el tiempo. La variable impulso se define de la siguiente forma:

$$I_t^{(h)} = \begin{cases} 0 & t \neq 0 \\ 1 & t = h \end{cases} \quad (3.18)$$

Si definimos el efecto del acontecimiento puntual como w_0 , la serie observada como z_t e y_t como la serie sin contaminar, entonces:

$$z_t = w_0 I_t^{(h)} + y_t = w_0 I_t^{(h)} + \psi(B)a_t \quad (3.19)$$

La expression general para incluir cualquier cantidad de eventos externos es la siguiente:

$$z_t = \sum_{i=0}^n w_i I_t^{(h+i)} + \psi(B)a_t \quad (3.20)$$

Es posible simplificar estas expresiones utilizando el operador de retardo, teniendo en cuenta que:

$$BI_t^{(h)} = I_{t-1}^{(h)} = I_t^{(h+1)} \quad (3.21)$$

Utilizando esta expresión, podemos definir la siguiente:

$$w(B) = w_0 + Bw_1 + B^2w_2 + \dots + B^n w_n \quad (3.22)$$

Finalmente, la expresión general compactada para un modelo con intervenciones basadas en variables impulso se muestra a continuación:

$$z_t = w(B)I_t^{(h)} + \psi(B)a_t \quad (3.23)$$

Las variables escalón sirven para construir intervenciones que representen los eventos que cambian de forma permanente el comportamiento de una serie temporal, es decir, cambios de nivel. Definimos la variable escalón de la siguiente forma:

$$S_t^{(h)} = \begin{cases} 0 & t < 0 \\ 1 & t \geq h \end{cases} \quad (3.24)$$

La expresión de la serie temporal en función del modelo ARIMA y esta intervención sería:

$$z_t = w_0 S_t^{(h)} + y_t = w_0 S_t^{(h)} + \psi(B)a_t \quad (3.25)$$

De esta forma, el efecto de cambio de nivel queda modelado en la serie temporal, ya que:

$$\begin{cases} z_t = y_t & t < h \\ z_{t+i} = w_0 + y_{h+i} & i \geq 0 \end{cases} \quad (3.26)$$

La expresión final para múltiples escalones, utilizando el operador de retardo es:

$$z_t = w(B)S_t^{(h)} + \psi(B)a_t \quad (3.27)$$

De esta forma, la serie temporal mostraría varios cambios de nivel tal y como se muestra a continuación:

$$z_t = y_t \quad t < h \quad (3.28)$$

$$z_h = w_0 + y_h$$

$$z_{h+1} = w_0 + w_1 + y_{h+1}$$

...

$$z_{h+i} = \sum_{i=0}^n w_i + y_{h+i}$$

3.13 Datos atípicos: outliers.

En el apartado anterior sobre intervenciones, se hablaba de eventos puntuales que modificaban el comportamiento de la serie, cuyos efectos podían ser modelados. Se daba por hecho que estos eventos eran conocidos y fácilmente identificables. Sin embargo, en la realidad muchas veces estos eventos no son conocidos, por lo que la serie que se está estudiando presenta comportamientos puntuales que no pueden ser explicados mediante un modelo ARIMA que si explica el resto de la serie. Estos comportamientos extraños se denominan valores atípicos o outliers, y su detección e identificación resulta de vital importancia para llevar a cabo un análisis exitoso de cualquier serie temporal.

Debido al estrecho paralelismo entre los valores atípicos y las intervenciones, es necesario establecer una comparación directa entre los dos para entender sus diferencias. Recordando el apartado anterior, el proceso para incluir una intervención en un modelo es el siguiente: se conoce un evento en el tiempo que afecta al comportamiento de la serie temporal en un dominio del tiempo concreto, pero se desconoce la magnitud de este efecto. Estudiando la naturaleza de dicho evento, se modela una intervención acorde. Esta intervención se incluye en el modelo como variable independiente explicativa, y se estima su valor (es decir, la magnitud del efecto sobre la serie) junto al resto de parámetros de la serie. Finalmente se estima un modelo ARIMA con sus parámetros, acompañado de la intervención y la magnitud de esta.

Con los valores atípicos, el procedimiento es el inverso: se dispone de una serie temporal, la cual, tras un proceso de análisis, se determina que puede ser explicada mediante un modelo ARIMA concreto. No obstante, se entiende que la existencia de eventos ajenos a nuestro conocimiento puede estar afectando al comportamiento de la serie. Su detección incluye conocer su valor, y por lo tanto la magnitud con la que afecta y modifica la serie, así como su naturaleza, permitiendo clasificarlo. Finalmente, se estima un modelo ARIMA final con sus parámetros, acompañado de los valores atípicos detectados, incluyendo su magnitud y naturaleza, y de las intervenciones incluidas en el modelo si las hubiere.

Atípicos aditivos (AO: additive outlier). Este tipo de valores atípicos son el equivalente a las intervenciones basadas en variables impulso. Se determina que hay presencia de un valor atípico aditivo si para un determinado instante h el valor de la serie temporal correspondiente a $t=h$ se ha generado de forma diferente al resto. Si definimos la magnitud de la diferencia

entre el valor esperado y el atípico como w_a , la serie temporal z_t bajo la influencia del AO queda expresada como:

$$z_t = \begin{cases} \psi(B)a_t & t \neq h \\ \psi(B)a_t + w_a & t = h \end{cases} \quad (3.29)$$

Usando la ya conocida definición de variable impulso, se puede escribir que:

$$z_t = w_a I_t^{(h)} + y_t = w_a I_t^{(h)} + \psi(B)a_t \quad (3.30)$$

La forma más compacta escribir los modelos afectados por AO es la siguiente:

$$(z_t - w_a I_t^{(h)}) = \psi(B)a_t \quad (3.31)$$

A continuación, se presenta un gráfico realizado mediante la función TSO en R, mostrando el efecto de un atípico aditivo sobre una serie temporal:

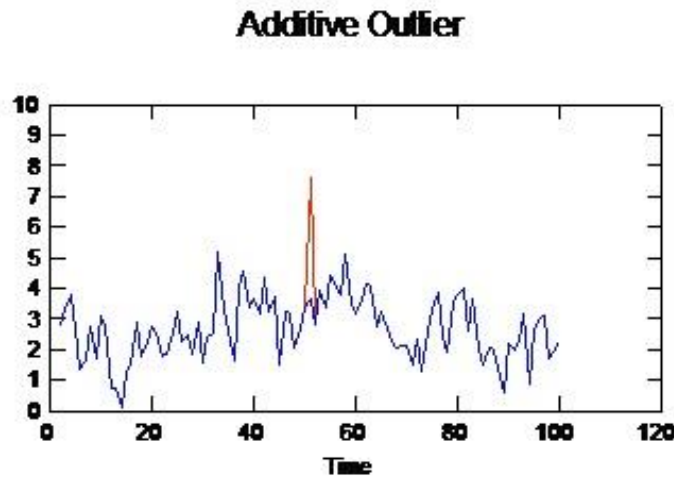


Figura 3.10. Efecto de un AO sobre una serie temporal (elaboración en R).

Atípicos innovativos (IO: innovative outliers). Este tipo de valores atípico afectan a la serie temporal actuando sobre la innovación en lugar de la observación en sí. Hay presencia de un IO en una serie temporal en el momento h si en dicho instante la innovación se ve afectada por un valor de magnitud desconocida.

$$z_t = \psi(B)(w_I I_t^{(h)} + a_t) \quad (3.32)$$

El efecto sobre la serie es el siguiente:

$$z_t = \begin{cases} \psi(B)a_t & t < h \\ \psi(B)a_t + w_I \psi_j & t = h + j, j \geq 0 \end{cases} \quad (3.33)$$

Los residuos para una serie temporal con presencia de un IO son:

$$e_t = w_I I_t^{(h)} + a_t = a_h + w_I \quad (3.34)$$

A continuación, se muestra el efecto de un atípico innovativo sobre una serie temporal:

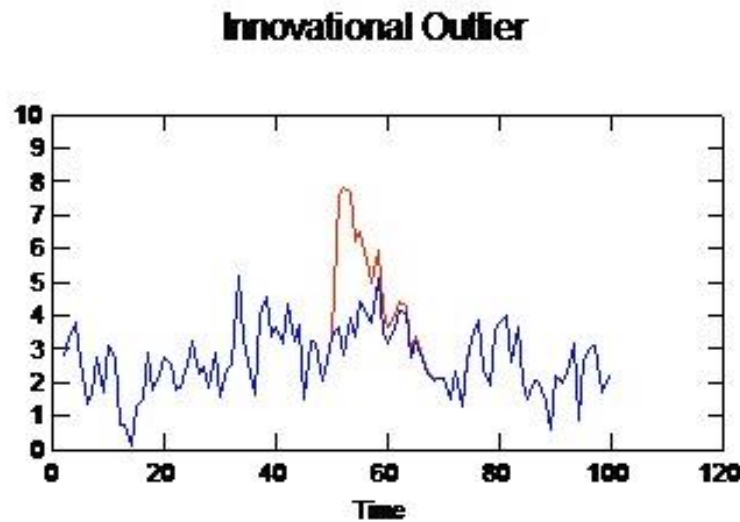


Figura 3.11. Efecto de un IO sobre una serie temporal (elaboración en R).

Cambios de nivel (LS: level shift). Estos tipos de valores atípicos son la contrapartida de las intervenciones basadas en variables escalón. Se caracterizan por provocar un cambio de nivel en la serie temporal a partir de instante h , dando lugar a:

$$z_t = w_L S_t^{(h)} + \psi(B) a_t \quad (3.35)$$

La relación con la serie sin contaminar es la siguiente:

$$z_t = \begin{cases} \psi(B)a_t & t < h \\ \psi(B)a_t + w_L & t \geq h + j \end{cases} \quad (3.36)$$

A continuación, se muestra el efecto de un atípico LS sobre una serie temporal:

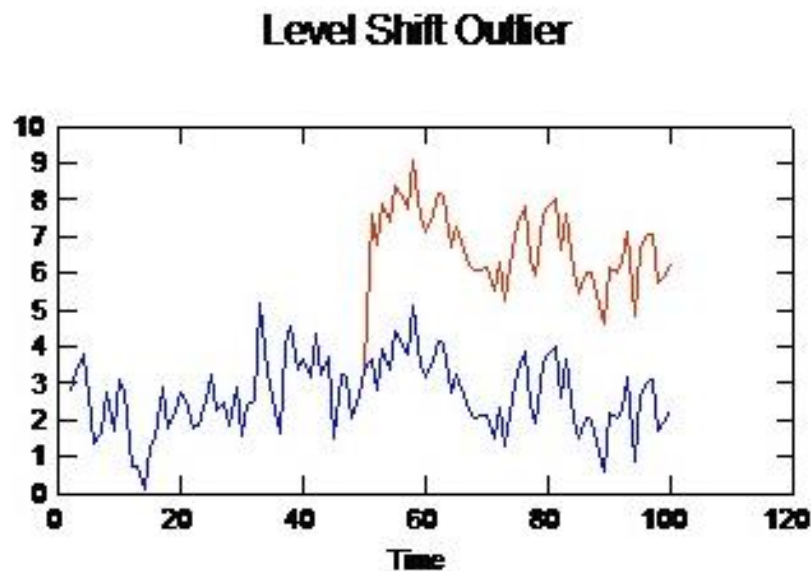


Figura 3.12. Efecto de un LS sobre una serie temporal (elaboración en R).

Cambios transitorios (TC). Este tipo de valores atípicos son más complejos. Su expresión es:

$$z_t = \frac{w_{TC}}{1-\delta B} I_t^{(h)} + \psi(B) a_t \quad (3.37)$$

Como se puede apreciar en la expresión matemática de este tipo de atípico, su naturaleza se encuentra en un punto intermedio entre un cambio de nivel y un atípico aditivo, y su comportamiento depende de valor de δ : si $\delta = 1$ se comportara como un cambio de nivel, y si $\delta = 0$ lo hará como un atípico aditivo. A continuación, se muestra el efecto de un atípico LS sobre una serie temporal:

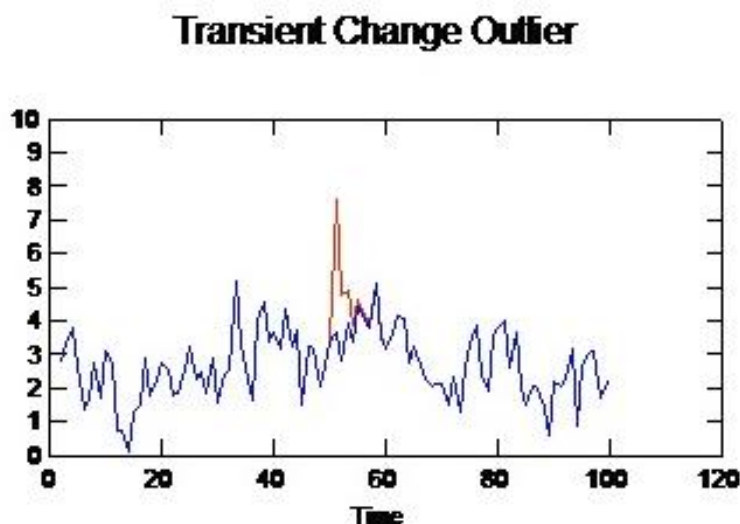


Figura 3.13. Efecto de un cambio transitorio sobre una serie temporal (elaboración en R).

Efecto rampa (R). Su efecto consiste en introducir una tendencia en la serie de pendiente w_R desde un determinado momento “h”.

3.14 Metodología de Box-Jenkins para el análisis de series temporales.

A lo largo de este trabajo, se utilizará el sistema Box-Jenkins para abordar el análisis de las series temporales a estudiar. Dicho método está basado en los conceptos teóricos expuestos en este capítulo, y consta de los siguientes pasos:

1. Análisis del gráfico de secuencia de la serie estudiada. Si se detecta que no es estacionaria, se llevan a cabo las transformaciones necesarias.
2. Obtención de la FAS y la FAP de la serie ya transformada en estacionaria. Se estudian ambas funciones de autocorrelación y se estima un modelo ARIMA a ajustar que explique el comportamiento de la serie.
3. Se procede a ajustar el modelo elegido con detección automática de outliers. Se comprueba la bondad del ajuste, estudiando los residuos generados por el modelo a través de su FAS y FAP, y se los somete a un test. Se comprueba que los parámetros del modelo resultan significativos. Se escribe la ecuación final del modelo ARIMA, y se exponen los resultados junto a la lista de atípicos detectados.

Si todos los pasos se han realizado correctamente y han sido superados con éxito, se pueden estimar predicciones basadas en el modelo ARIMA obtenido.

4 SERIES ANALIZADAS.

En este capítulo se procederá a analizar series temporales pertenecientes a diferentes sectores: sector de la construcción, sector inmobiliario, industria metalúrgica e indicadores económicos generales. Para ello, se hará uso de la teoría descrita en el apartado anterior, así como del método Box-Jenkins. Los ajustes de los modelos se realizarán en R y SPSS. Las primeras series se mostrarán con detalle, con el fin de explicar los pasos necesarios para analizar una serie. El resto de las series se presentarán de forma resumida.

A continuación, se muestra una tabla recopilando todas las series analizadas, así como la información esencial sobre ellas. Se incluye una codificación numérica que se utilizará a lo largo del trabajo para referirse de forma rápida a las series sin necesidad de escribir el nombre completo de estas.

Nombre y código	Fuente	Periodicidad Inicio - Fin	Información
S1- Trabajadores asociados a la seguridad social en sector de la construcción	Banco de España – Ministerio de empleo y seguridad social	Mensual 01/1995 -09/2017	Uds.: personas
S2- Índice de precios de la mano de obra en el sector de la construcción	Comité Superior de Precios de Contratos del Estado (CSPCE)	Mensual 01/1995- 12/2016	Calculado a partir de las fórmulas incluidas en el anexo
S3- Parados sector construcción	Banco de España - INE	Mensual 01/1996- 09/2017	Uds.: Personas

S4- Asalariados en el sector de la construcción	Banco de España - INE	Trimestral 03/2005 (T1) -06/2017 (T2)	Uds.: Miles de personas
S5- Número de compraventas. Fincas de naturaleza urbana: viviendas. Segunda mano.	Banco de España- INE	Mensual 03/2007 -08/2017	Uds.: Número de operaciones
S6- Número de compraventas. Fincas de naturaleza urbana: viviendas. Nueva.	Banco de España- INE	Mensual 03/2007 -08/2017	Uds.: Número de operaciones
S7- Hipotecas totales constituidas en España.	INE	Mensual 01/2003- 09/2017	Uds.: número de créditos concedidos.
S8- Índice IPI industria metalúrgica	Banco de España - INE	Mensual 01/1975 -08/2017	Uds.: Exponente -2
S9- Producción interna de acero	Banco de España - Unesid	Mensual 01/1967 -05/2017	Uds.: Miles de toneladas métricas
S10- Importaciones acero	Banco de España - Unesid	Mensual 01/1967 -05/2017	Uds.: Miles de toneladas métricas
S11- Exportaciones acero	Banco de España - Unesid	Mensual 01/1967 -05/2017	Uds.: Miles de toneladas métricas

S12- Paro total registrado en España	INE	Mensual	Uds.: personas
S13-- IPC nacional	INE	Mensual 01/2002- 09/2017	-

Tabla 4.1. Listado de series temporales a analizar.

4.1 Trabajadores asociados a la seguridad social en el sector de la construcción.

Los datos obtenidos para esta serie temporal van desde enero de 1995 hasta septiembre de 2017. A continuación, se muestra su gráfico de secuencia de la serie, seguida de su FAS y FAP:

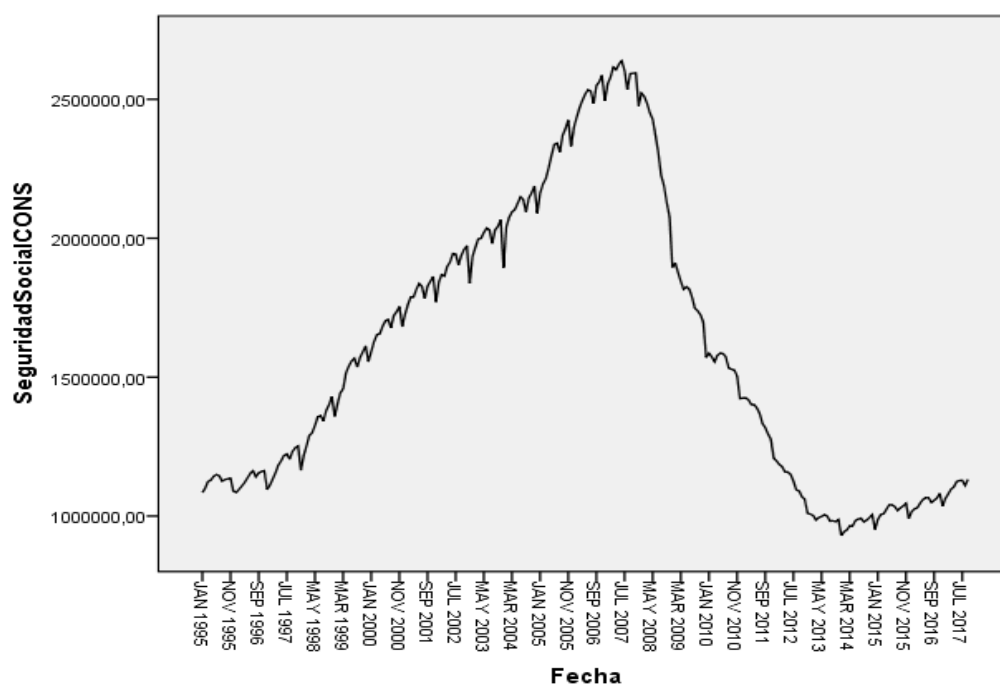


Figura 4.1. Gráfico de secuencia para la serie temporal número de asociados a la seguridad social en el sector de la construcción.

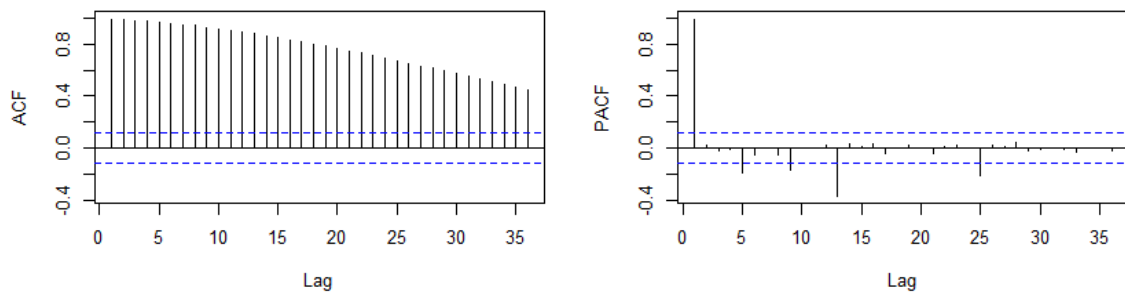


Figura 4.2. FAS y FAP del número de asociados a la seguridad social en el sector de la construcción.

Tomamos una diferencia regular y otra estacional con el fin de transformar la serie en estacionaria. Tras estas transformaciones, el P valor del test de Dickey-Fuller es menor que 0.01, por lo que la serie se puede considerar estacionaria. La imagen inferior muestra la serie ya transformada, junto a su FAS y FAP.

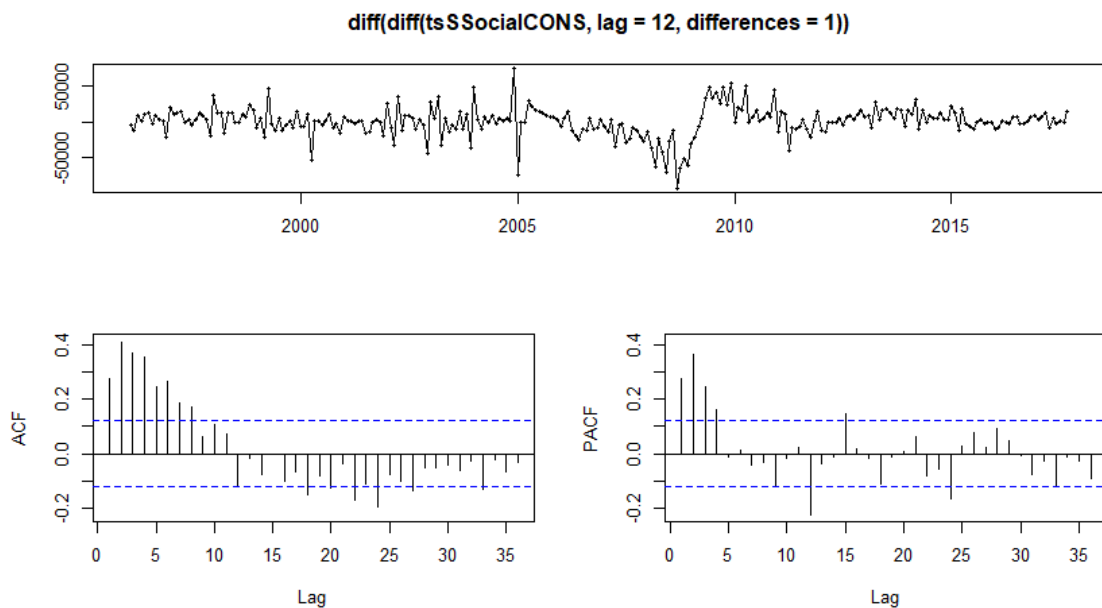


Figura 4.3. Gráfico de secuencia, FAS y FAP de la serie para el número de asociados a la seguridad social en el sector de la construcción transformada en estacionaria.

Observando la FAS y la FAP de la serie transformada en estacionaria, se propone un modelo ARIMA (1, 1, 1) x (2, 1, 0). Tras el ajuste, la ecuación que explica el comportamiento de la serie es la siguiente:

$$(1 + 0.293B^{12} + 0.315B^{24})(1 - 0.906B)\nabla\nabla_{12}z_t = (1 - 0.732B)a_t \quad (4.1)$$

$$\begin{matrix} (-4.329) & (-4.792) & (42.527) & (12.124) \end{matrix}$$

Observando las FAS y FAP de los residuos generados por el modelo, se observa que ninguno de los palos supera los límites de las bandas de confianza:

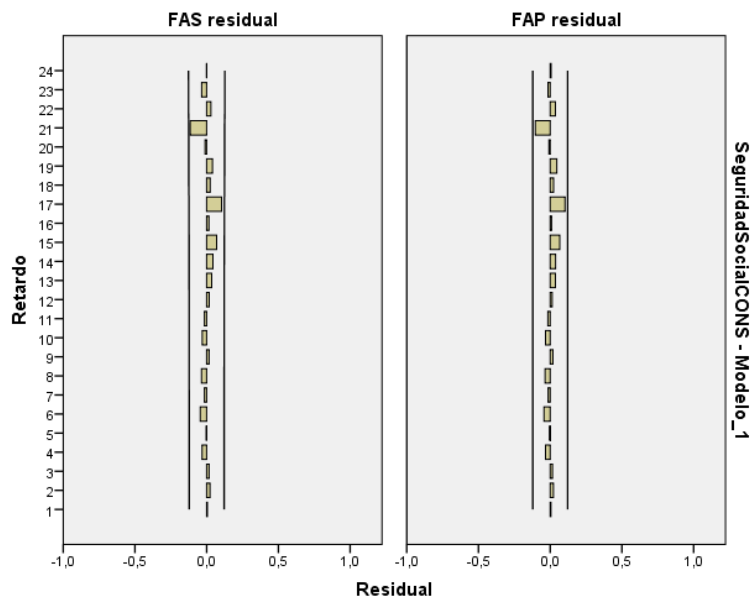


Figura 4.4. FAS y FAP residual para el ajuste de la serie número de asociados a la seguridad social en el sector de la construcción.

Además, los residuos pasan el test de Ljung-Box con un P-valor de 0.919. Teniendo en cuenta que además todos los parámetros obtenidos para el modelo son significativos, podemos considerar el ajuste como bueno. A continuación, se muestra una tabla con todos los atípicos detectados para esta serie, así como un gráfico mostrando el efecto de estos atípicos sobre la serie.

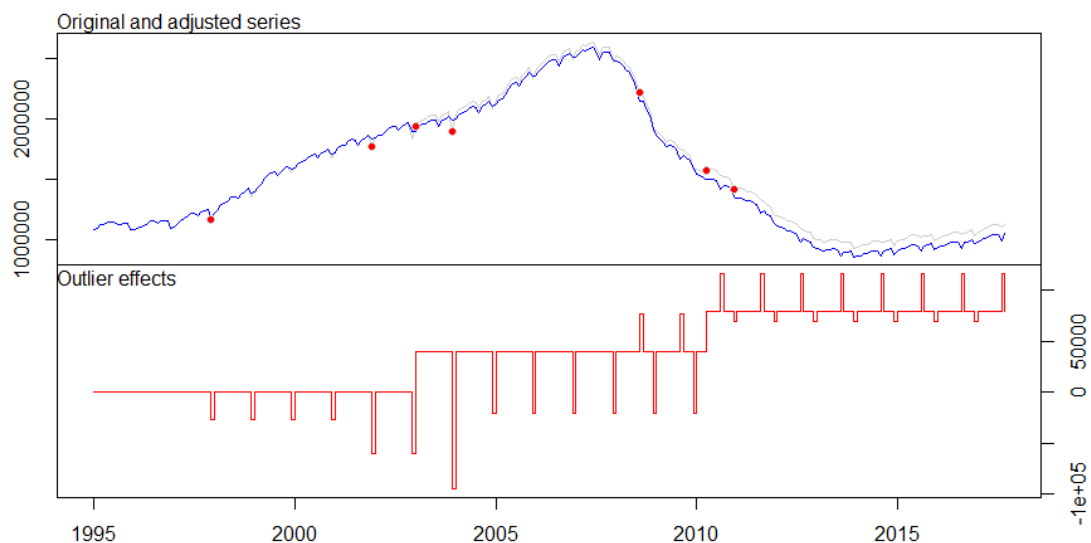


Figura 4.5. Impacto de los outliers en la serie temporal número de asociados a la seguridad social en el sector de la construcción.

Valores atípicos

			Estimación	SE	t	Sig.
SeguridadSocialCONS-Modelo_1	Dic 1997	Aditivo estacional	-26657,954	5296,042	-5,034	,000
	Abr 1999	Cambio de nivel	51696,917	6726,791	7,685	,000
	Dic 2001	Aditivo estacional	-23505,847	5311,202	-4,426	,000
	Mar 2002	Aditivo	-30978,540	4413,945	-7,018	,000
	Dic 2002	Aditivo	-38752,224	4987,185	-7,770	,000
	Dic 2003	Aditivo	-78031,087	5095,402	-15,314	,000
	Abr 2005	Tendencia local	18924,861	4786,131	3,954	,000
	Mar 2007	Aditivo	18047,417	4410,687	4,092	,000
	Jun 2007	Aditivo	17762,107	4341,648	4,091	,000
	Feb 2008	Tendencia local	-35779,075	4991,585	-7,168	,000
	Ago 2008	Aditivo estacional	36675,511	5071,588	7,232	,000
	Dic 2008	Aditivo	-22467,866	4355,480	-5,159	,000
	Mayo 2009	Tendencia local	29476,782	5221,515	5,645	,000

Abr 2010	Cambio de nivel	33171,079	6990,501	4,745	,000
Dic 2010	Aditivo estacional	38041,761	5067,418	7,507	,000
Mar 2013	Aditivo	-17779,841	4396,150	-4,044	,000

Tabla 4.2. Atípicos detectados en SPSS para la serie temporal “número de asociados a la seguridad social en el sector de la construcción”.

4.2 Índice nacional del precio de la mano de obra en el sector de la construcción en España.

El índice nacional del precio de la mano de obra en el sector de la construcción es un índice utilizado para la medición del coste de contratación de la mano de obra en proyectos de construcción de diversa índole de cara a la administración pública. Los datos para esta serie temporal comienzan en enero de 2002 y terminan en diciembre de 2016. El gráfico de secuencia correspondiente se muestra a continuación, junto a su FAS y FAP:

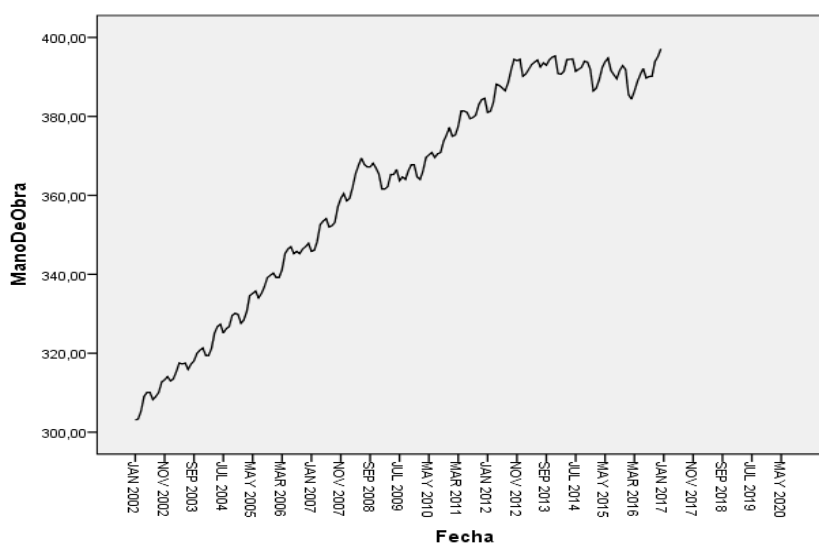


Figura 4.6. Gráfico de secuencia para la serie de índice de precios de la mano de obra en el sector de la construcción.

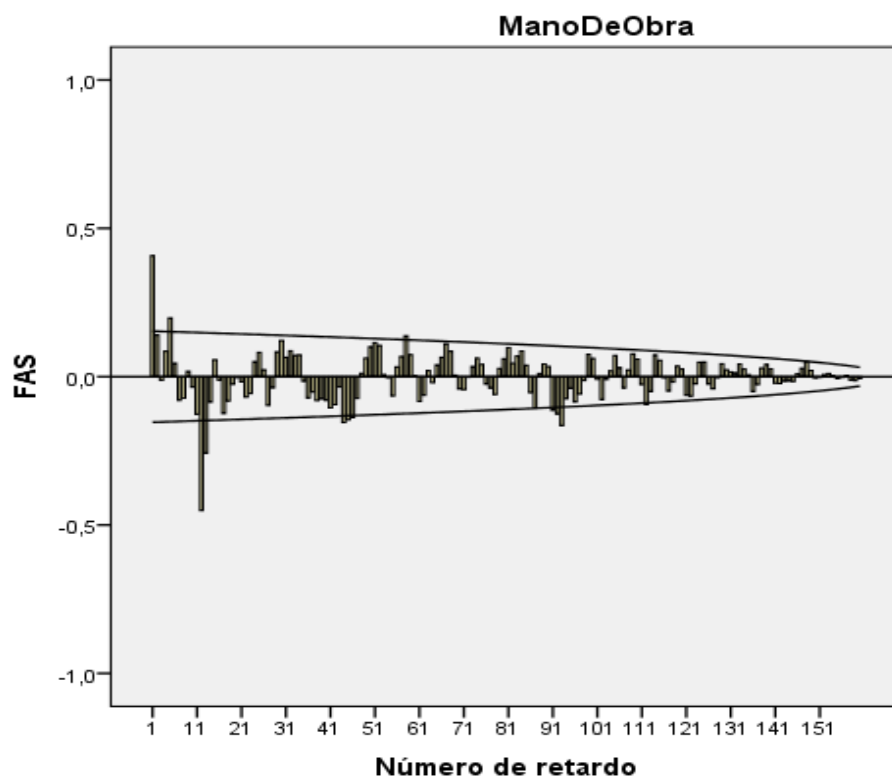


Figura 4.7. FAS para la serie del índice de precios de la mano de obra en el sector de la construcción.

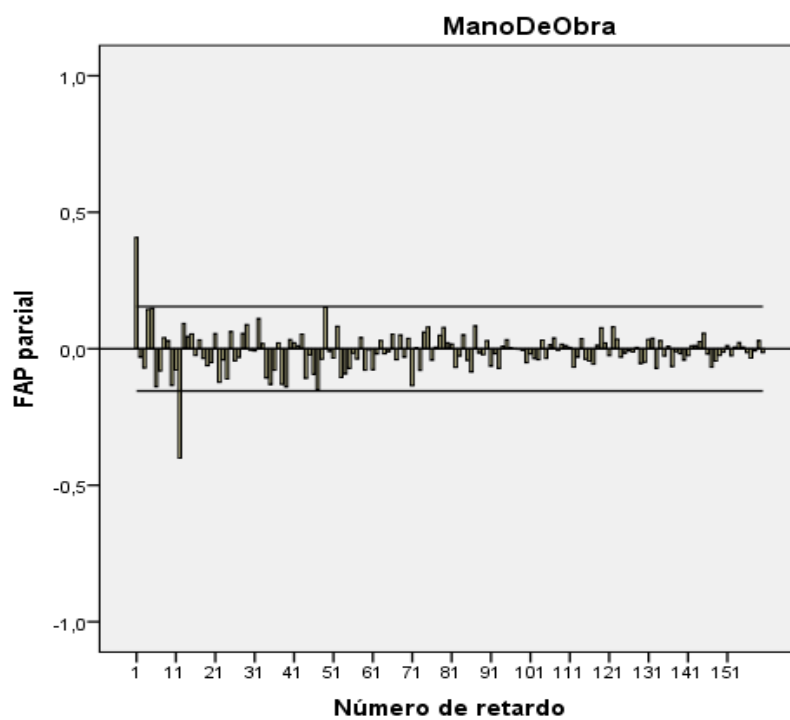


Figura 4.8. FAP para la serie del índice de precios de la mano de obra en el sector de la construcción.

Procedemos a analizar la serie temporal por el mismo proceso que la serie anterior. Observamos claramente que la serie no es estacionaria. Se toma una diferencia regular y otra estacional para transformar la serie en estacionaria. Los resultados de la transformación se muestran a continuación:

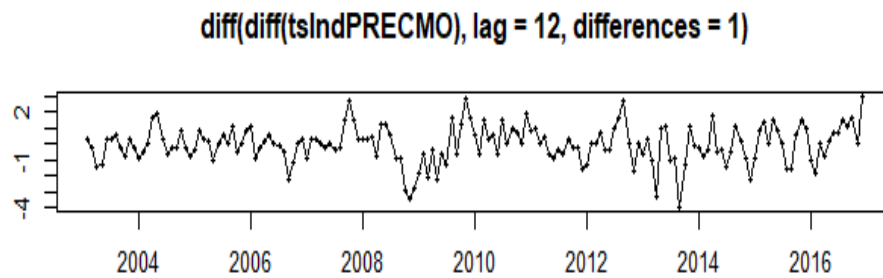


Figura 4.9. Serie del índice de precios de la mano de obra en el sector de la construcción transformada en estacionaria.

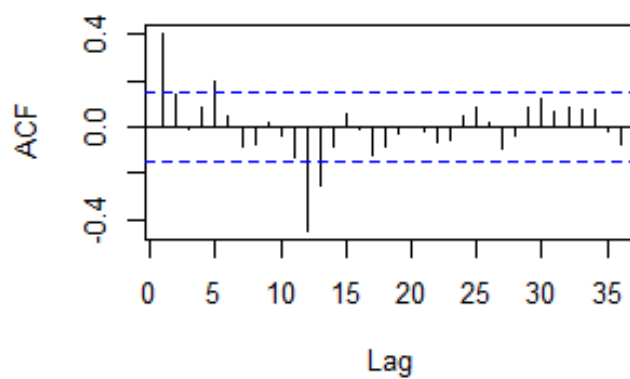


Figura 4.10. FAS del índice de precios de la mano de obra en el sector de la construcción transformada en estacionaria.

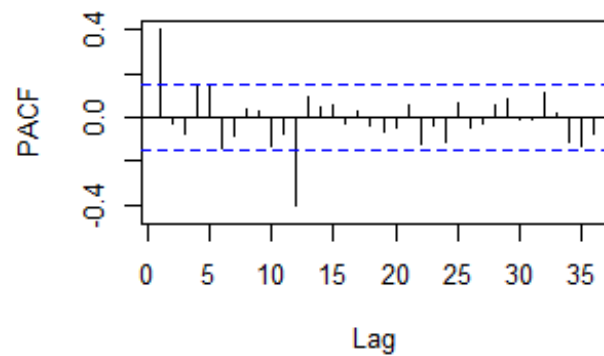


Figura 4.11. FAP del índice de precios de la mano de obra en el sector de la construcción transformada en estacionaria.

Una vez que la serie ha sido transformada en estacionaria, procedemos a estudiar la estructura de sus funciones de autocorrelación. Se propone ajustar un modelo ARIMA (1,1,0) x (0,1,1). La ecuación del modelo ajustado es la siguiente:

$$(1 - 0.4216B)\nabla\nabla_{12}z_t = (1 - 0.6980B^{12})a_t \quad (4.2)$$

(5.831) (10.49)

La FAS y la FAP de los residuos generados en el ajuste no muestran palos que superen las bandas de confianza tal y como muestra la siguiente imagen:

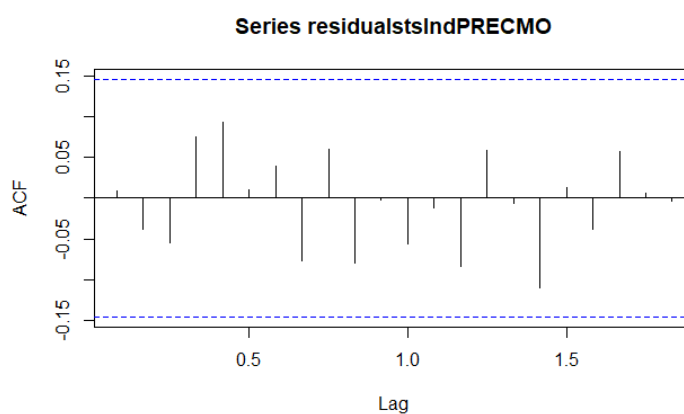


Figura 4.12. FAS de los residuos para el índice de precios de la mano de obra en el sector de la construcción.

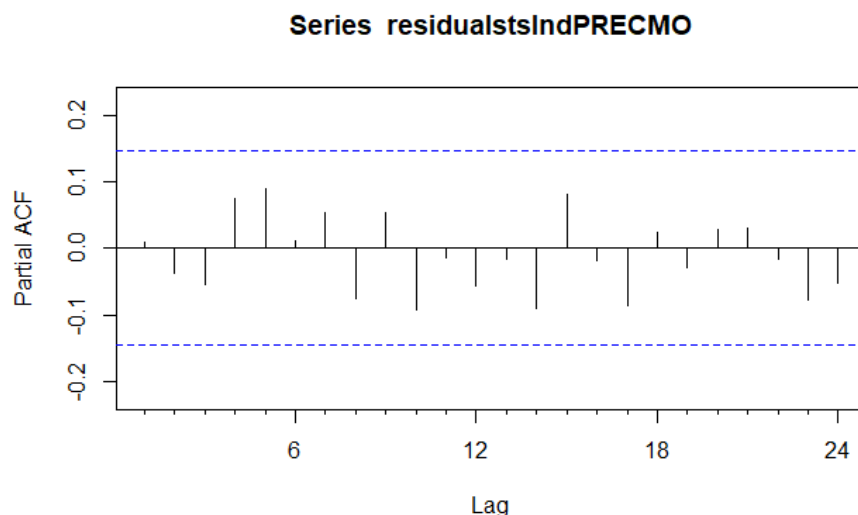


Figura 4.13. FAP de los residuos para el índice de precios de la mano de obra en el sector de la construcción.

Los residuos superan los test de Ljung-Box y Box-pierce con respectivos P-valores de 0.9052 y 0.906. A raíz del estudio de los residuos, y de la significancia de todos parámetros del modelo, se considera que el modelo elegido es correcto. Ya que esta serie ha sido analizada en R, se utiliza la función “auto.arima” para corroborar que el modelo elegido es el correcto. Tal y como se muestra a continuación, la función considera que el mejor modelo a ajustar es el mismo que hemos elegido. (Extracto de los resultados en la consola de R).

```
Series: tsIndPRECMO
ARIMA(1,1,0)(0,1,1)[12]

Coefficients:
      ar1      sma1
    0.4269  -0.7168
s.e.    0.0714   0.0669

sigma^2 estimated as 0.744:  log likelihood=-214.9
AIC=435.81  AICC=435.95  BIC=445.16
```

Los atípicos detectados para esta serie son los siguientes:

Tipo	Posición	Coeficiente	Sig.
Cambio de nivel (LS)	129 – septiembre 2012	2.564	3.83

Tabla 4.3. Valores atípicos detectados para la serie de índice de precios de mano de obra en el sector de la construcción.

A continuación, se muestra el efecto de este atípico sobre la serie:

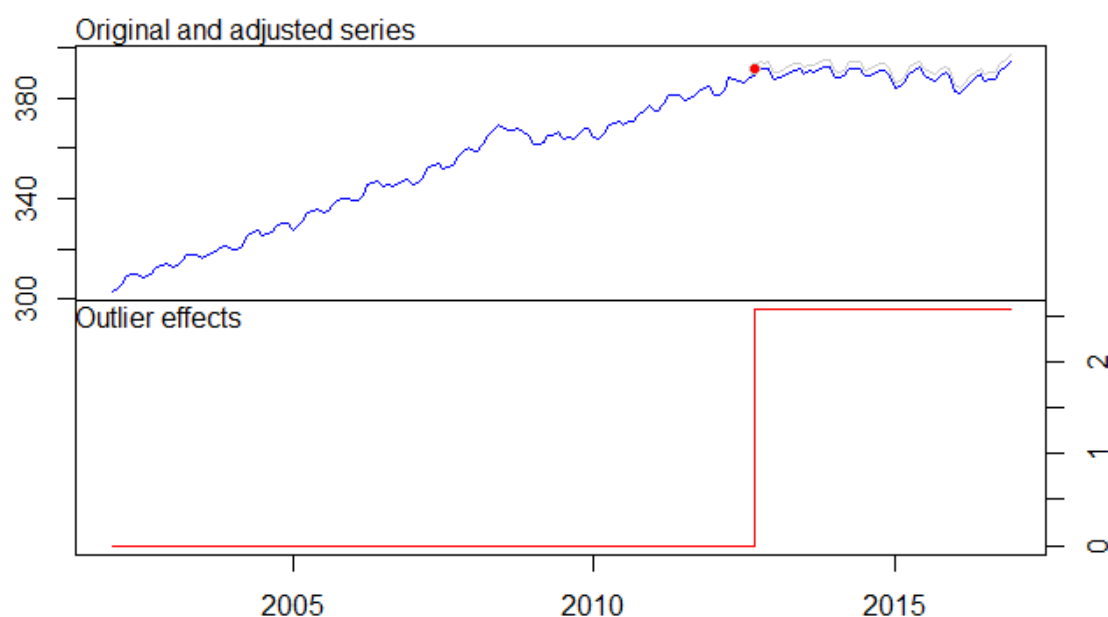


Figura 4.14. Gráfica de la serie índice de precios de la mano de obra en el sector de la construcción con los efectos de los outliers detectados.

4.3 Parados en el sector de la construcción.

Esta serie temporal representa el número total de parados, se ambos sexos, en el sector de la construcción. Cuenta con datos desde enero de 1996 hasta septiembre de 2017. El gráfico de secuencia de la serie temporal para los parados en el sector de la construcción se muestra a continuación:

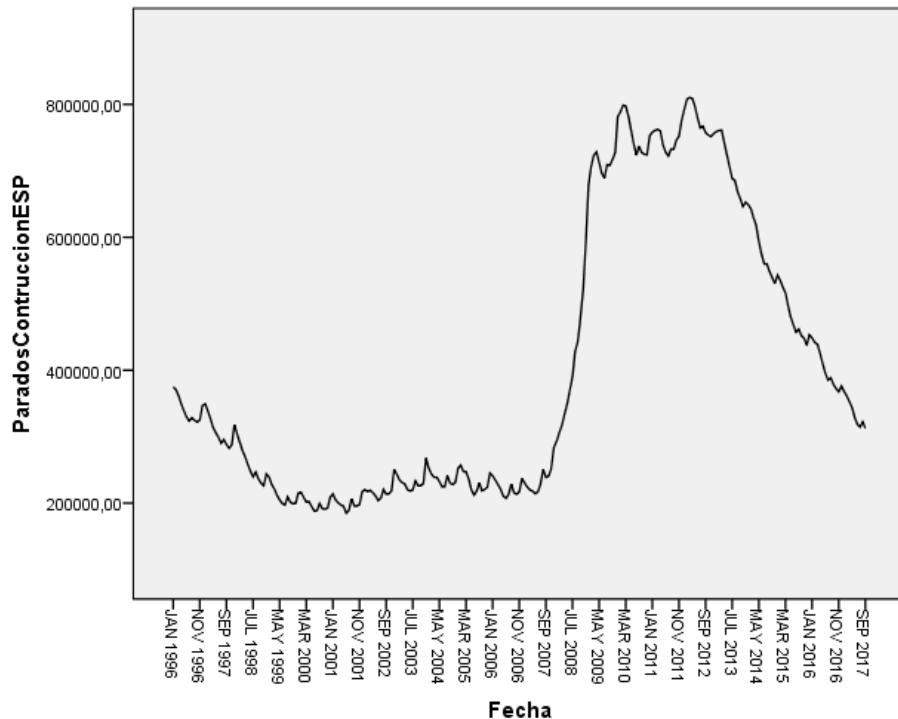


Figura 4.15. Gráfico de secuencia de los parados en el sector e la construcción en España.

Una primera inspección visual nos permite determinar que no es una serie estacionaria:

- La serie posee una clara tendencia descendente inicial. En torno a los años 2007 y 2008 presentar una agresiva tendencia ascendente, que vuelve a cambiar a partir de finales de 2013, volviendo a tener una fuerte tendencia descendente. Este brusco cambio a partir de 2007 es debido a la crisis económica.
- La serie presenta ciclo estacional mensual. Se observa un pico anual en el número de parados entorno a los meses de diciembre, enero y febrero, para comenzar a descender a partir de febrero hasta el siguiente diciembre, salvo por un pequeño y puntual remonte en los meses de agosto.
- La serie temporal es claramente heterocedástica.

Para obtener más información sobre el comportamiento de la serie y hacer predicciones fiables es necesario que procedamos a sus análisis con ayuda de R y SPSS. Gracias a la inspección visual realizada ya sabemos que la serie temporal no es estacionaria, aun así, se procede a descomponerla en series aditivas:

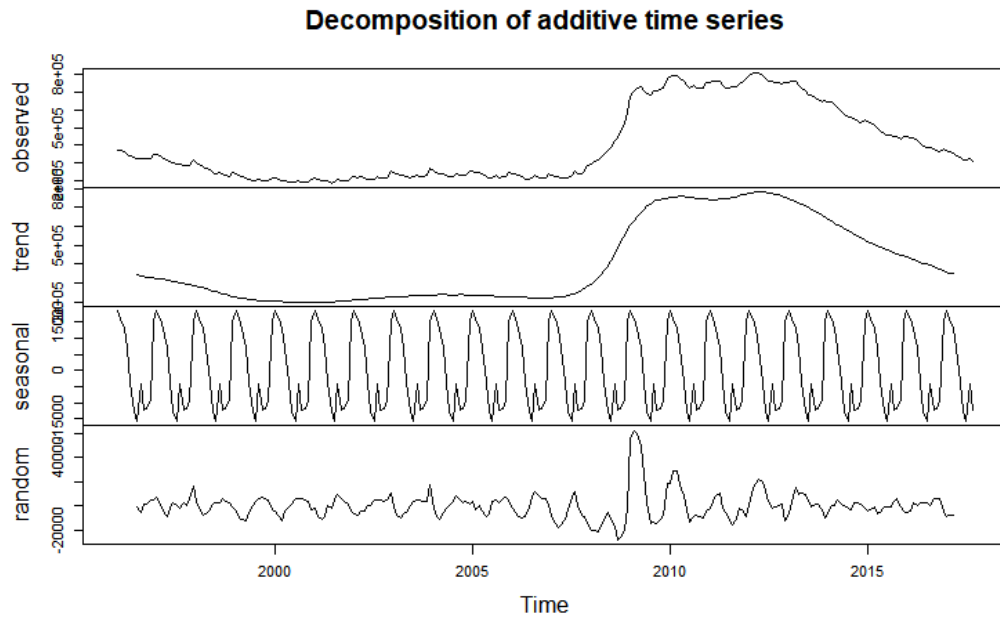


Figura 4.16. Descomposición de la serie temporal "parados en el sector de la construcción" en series aditivas de tendencia, estacionalidad y valores aleatorios.

La descomposición confirma nuestras hipótesis, mostrando claramente la tendencia de la serie y la componente estacional cíclica que se había observado. Siguiendo el proceso de análisis, se estudia la FAS y la FAP de la serie.

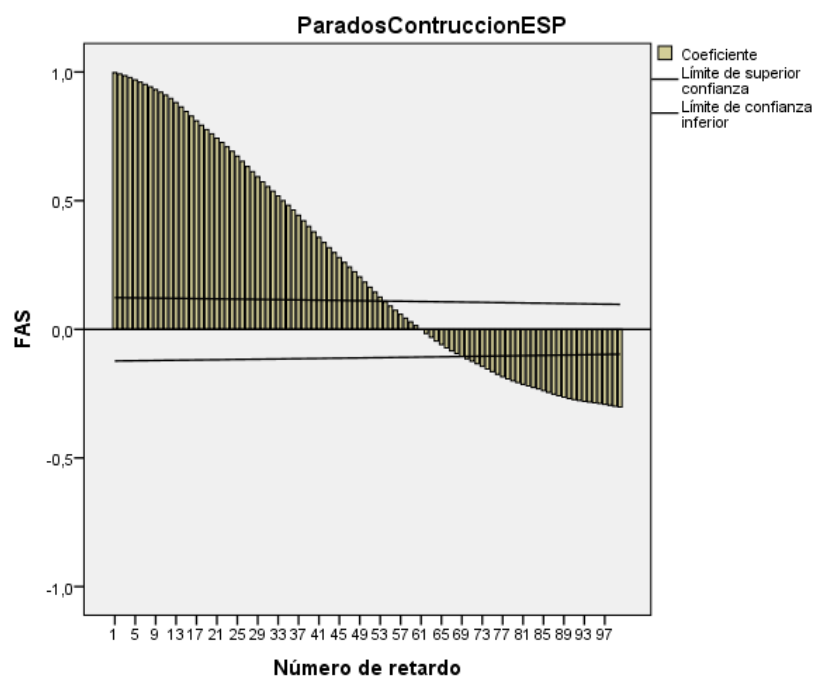


Figura 4.17. FAS parados en el sector de la construcción.

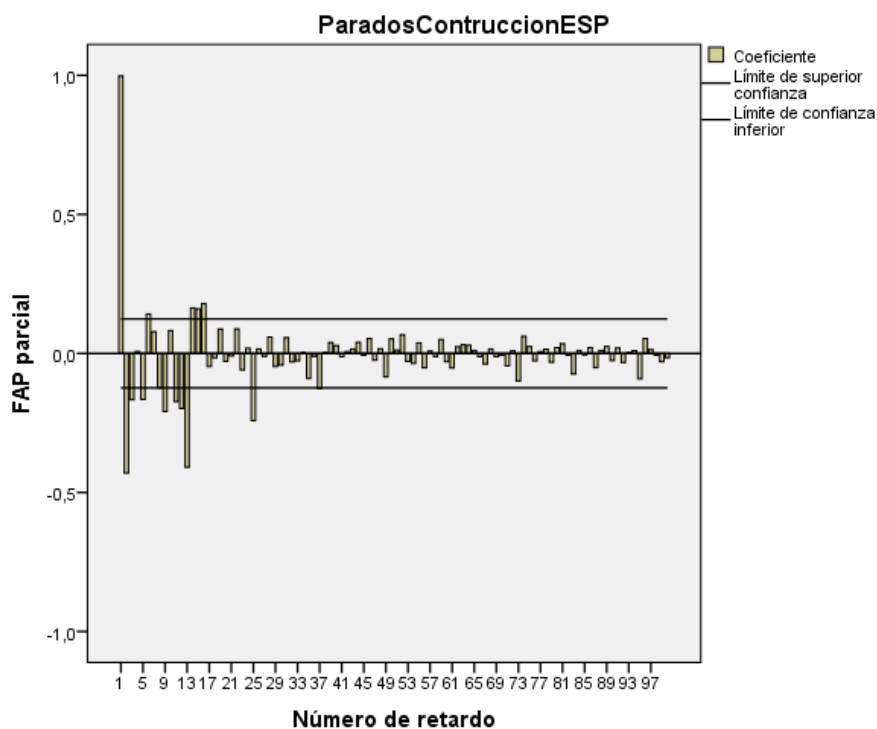


Figura 4.18. FAP parados en el sector de la construcción.

Observando la FAS, se puede ver que los palos no tienen a cero, por lo que la serie no es estacionaria. Para poder trabajar con la serie, necesitamos que sea estacionaria, por lo que se llevan a cabo varias transformaciones con tal fin. En primer lugar, tomamos una diferencia estacional para eliminar el ciclo mensual que presenta la serie. A continuación, se muestran los efectos de la transformación en la serie.

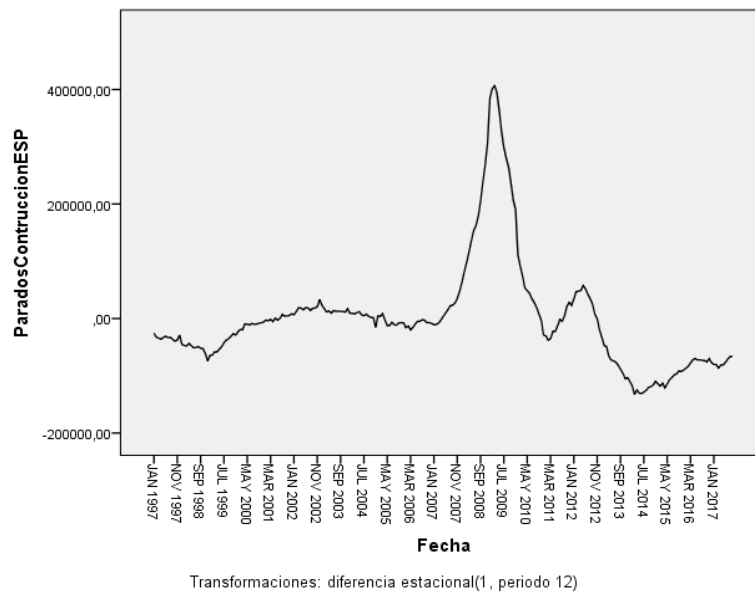


Figura 4.19. Serie temporal "parados en el sector de la construcción" tras tomar una diferencia estacional.

Tras tomar la diferencia estacional, observamos que la serie ya no presenta ciclo mensual de ningún tipo. Para eliminar la tendencia, tomamos una diferencia regular. La serie de parados en el sector de la construcción tras esta segunda transformación se muestra en la imagen inferior.

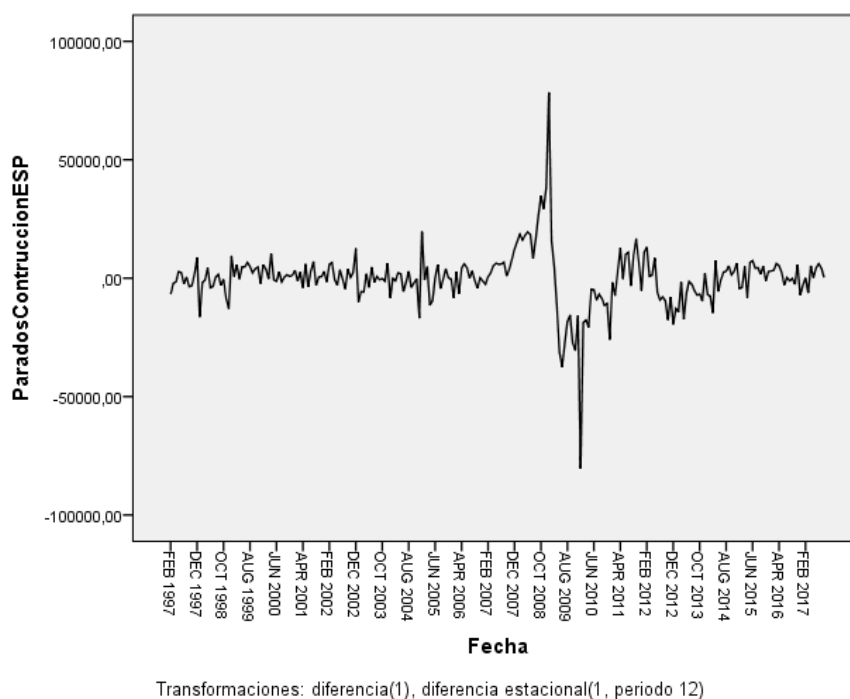
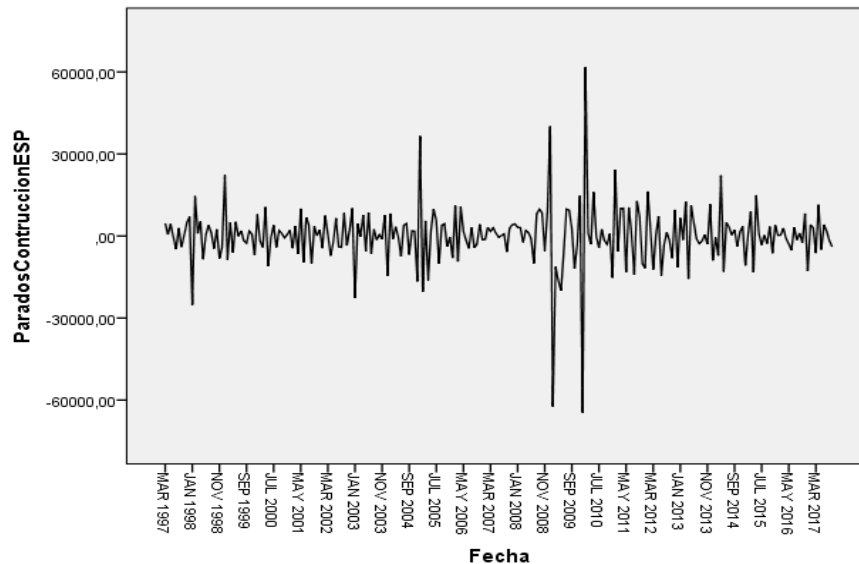


Figura 4.20. Serie temporal "parados en el sector de la construcción" con una diferencia regular y una diferencia estacional.

Para terminar de transformar la serie en estacionaria, tomamos una segunda diferencia. La serie tras llevarse a cabo todas las transformaciones mencionados presenta el siguiente aspecto:



Transformaciones: diferencia(2), diferencia estacional(1, periodo 12)

Figura 4.21. Serie de los parados en el sector de la construcción transformada en estacionaria tras aplicar una diferencia estacional, una diferencia estacional y tomar el logaritmo natural de la serie.

Las FAS y FAP para la serie transformada en estacionaria se muestran a continuación.

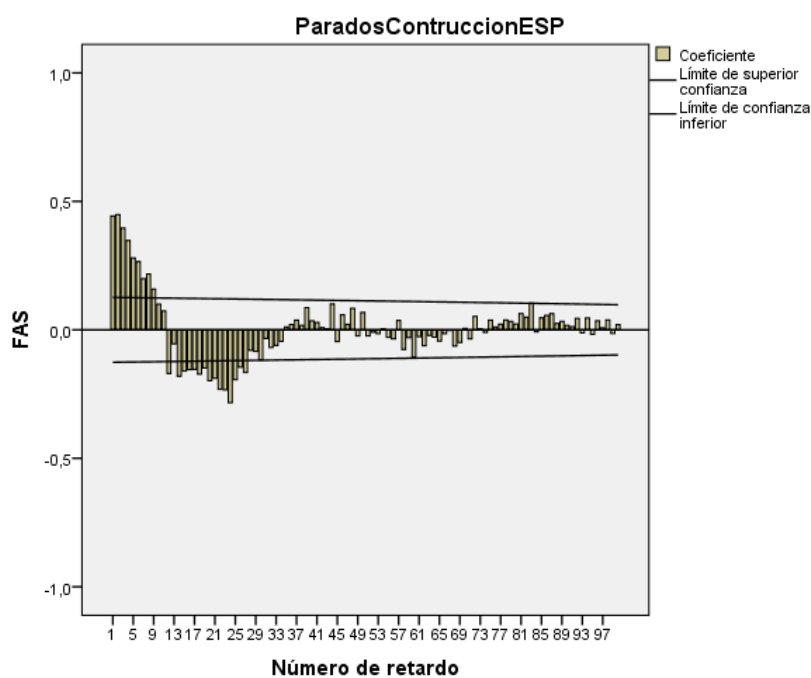


Figura 4.22. FAS parados en el sector de la construcción (estacionaria).

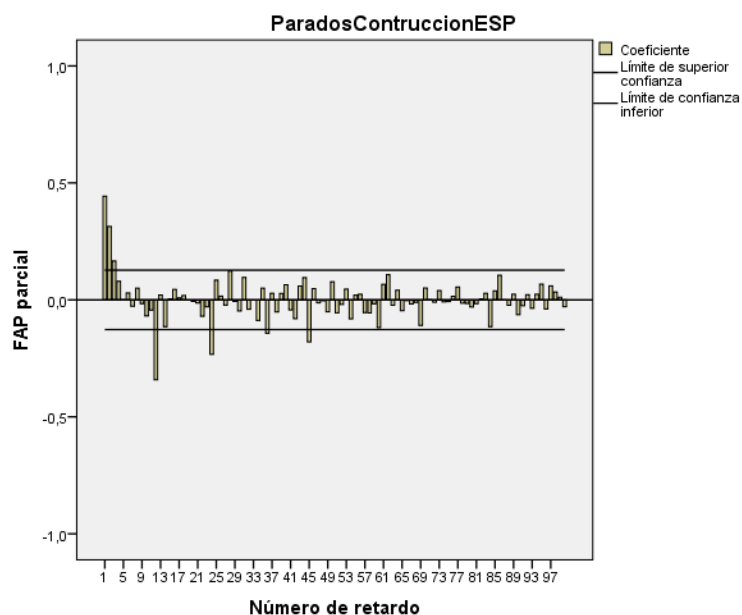


Figura 4.23. FAP parados en el sector de la construcción (estacionaria).

Tras estas transformaciones, la serie supera el test de Dickey-Fuller con un P-valor menor que 0.01, por lo que puede considerarse estacionaria. Estudiando las funciones de autocorrelación obtenidas, se decide aplicar a la serie un modelo ARIMA (0, 2, 1) x (2, 1, 0). Procedemos a estudiar los residuos del modelo ARIMA propuesto. A continuación, se muestra las FAS y la FAP de dichos residuos:

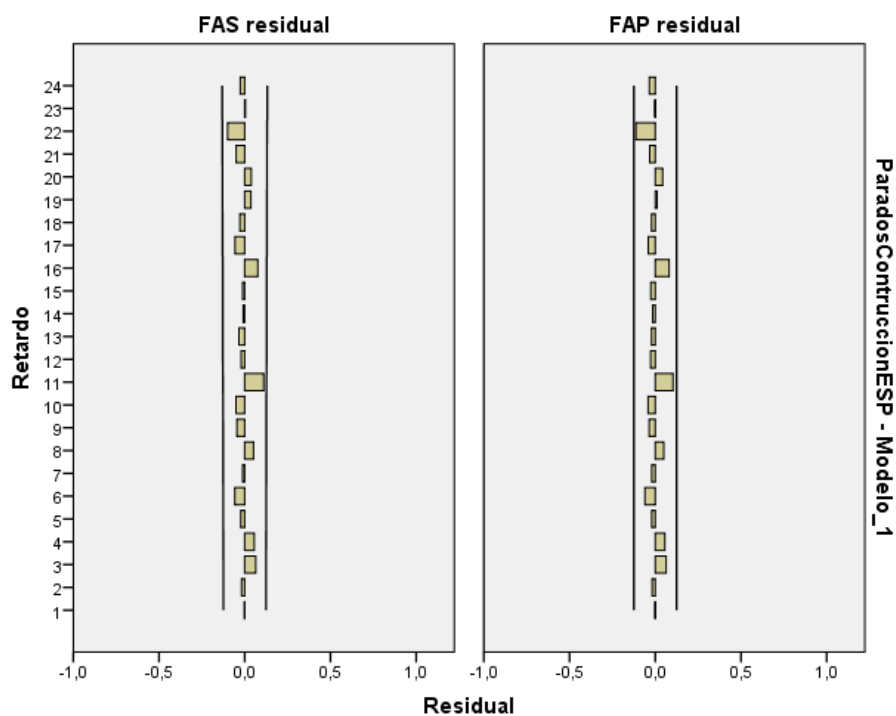


Figura 4.24. Funciones de autocorrelación para el modelo ARIMA propuesto para la serie de parados en el sector de la construcción.

No se aprecia ningún palo significativo que supere las bandas de confianza en las funciones de autocorrelación. Esto sugiere que los residuos obtenidos para el modelo ajustado son ruido blanco. Si los residuos del modelo son ruido blanco, significa que no muestran ningún tipo de relación entre ellos, por lo que a priori el ajuste puede ser bueno.

Continuamos el análisis estudiando las estimaciones arrojadas para los parámetros del modelo ARIMA, que son los siguientes:

Parámetros del modelo ARIMA				Estimación	SE	t	Sig.
ParadosContruccionESP-Modelo_1	ParadosContruccionESP	Sin transformación	Diferencia	2			
			MA Retardo 1	,466	,058	7,990	,000
			AR, estacional Retardo 1	-,305	,064	-4,756	,000
			Retardo 2	-,246	,062	-3,964	,000
			Diferencia estacional	1			

Tabla 4.4. Parámetros modelo ARIMA para la serie de parados en el sector de la construcción.

Observamos que todos los parámetros del modelo resultan significativos, ya que el valor absoluto de t para todos ellos es mayor que dos. Para finalizar el análisis, observamos el resultado del test de Ljung-Box.

Estadísticos del modelo						
Modelo	Número de predictores	Estadísticos de ajuste del modelo R cuadrado estacionaria	Ljung-Box Q(18)			Número de valores atípicos
			Estadísticos	QL	Sig.	
ParadosContruccionESP-Modelo_1	0	,780	11,105	14	,678	12

Tabla 4.5. Resultados generales para el ajuste del modelo ARIMA propuesto para la serie temporal de parados en el sector de la construcción.

El resultado del test es positivo. El estadístico de Ljung-Box constata que la afirmación de que las primeras 18 autocorrelaciones de los residuos son 0, tiene un nivel de significancia de 0.679, por lo que el modelo es válido. Podemos pasar a escribir la expresión final que rige el comportamiento de la serie temporal que hemos analizado:

$$(1 + 0.305B^{12} + 0.246B^{24})\nabla^2 \ln(z_t) = (1 - 0.466B)a_t \quad (4.3)$$

(-4.756)

(-3.964)

(7.990)

Para el modelo ARIMA ajustado, los siguientes atípicos han sido detectados:

Valores atípicos				Estimación	SE	t	Sig.
ParadosConstruccionESP-Modelo_1	Dic 1997	Aditivo		,034	,006	5,631	,000
	Abr 2000	Aditivo		,026	,006	4,267	,000
	Ago 2001	Aditivo		,025	,006	4,087	,000
	Dic 2002	Aditivo		,041	,007	5,731	,000
	Dic 2003	Aditivo		,069	,007	9,698	,000
	Mar 2005	Aditivo		,026	,006	4,245	,000
	Mayo 2005	Cambio de nivel		-,035	,009	-3,798	,000
	Dic 2007	Tendencia local		,042	,007	6,179	,000
	Jul 2008	Innovador		-,050	,013	-3,881	,000
	Ago 2008	Aditivo estacional		-,039	,008	-4,907	,000
	Ene 2009	Transitorio	Magnitud	,090	,010	9,428	,000
			Factor de decrecimiento	,780	,115	6,765	,000
	Mayo 2009	Tendencia local		-,042	,008	-5,470	,000

Tabla 4.6. Valores atípicos detectados para la serie de parados en el sector de la construcción.

A continuación, se muestra un gráfico de la serie afectada por los outliers detectados.

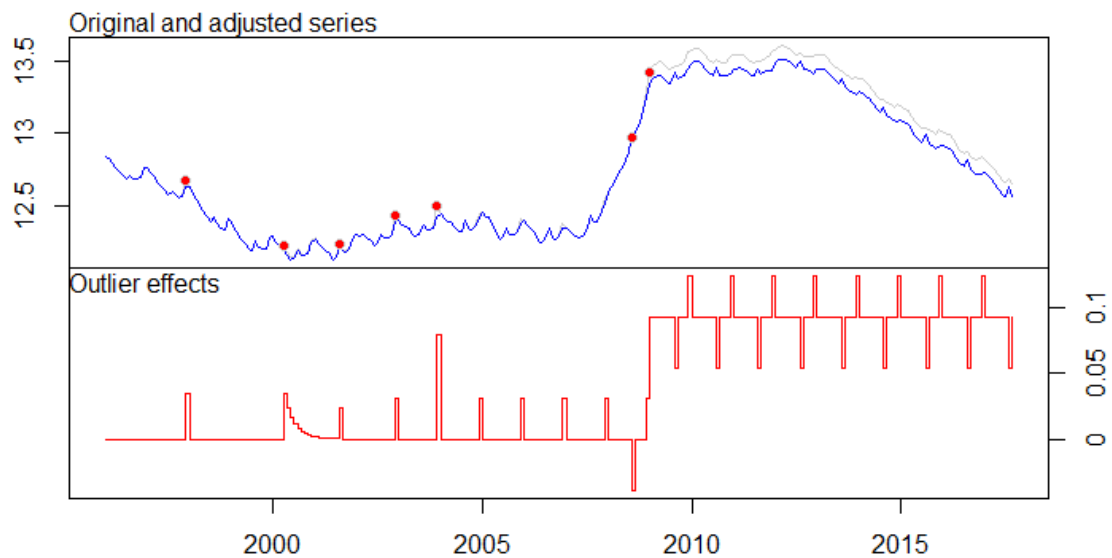


Figura 4.25. Gráfica de la serie "parados del sector de la construcción" con los efectos de los outliers detectados.

4.4 Número de asalariados en el sector de la construcción.

A diferencia de las series analizadas hasta el momento, esta serie es de carácter trimestral. Los datos van desde el primer trimestre de 2005 hasta el segundo trimestre de 2017. Las unidades son “miles de personas”. A continuación, se muestra el gráfico de secuencia de la serie junto a su FAS y FAP.

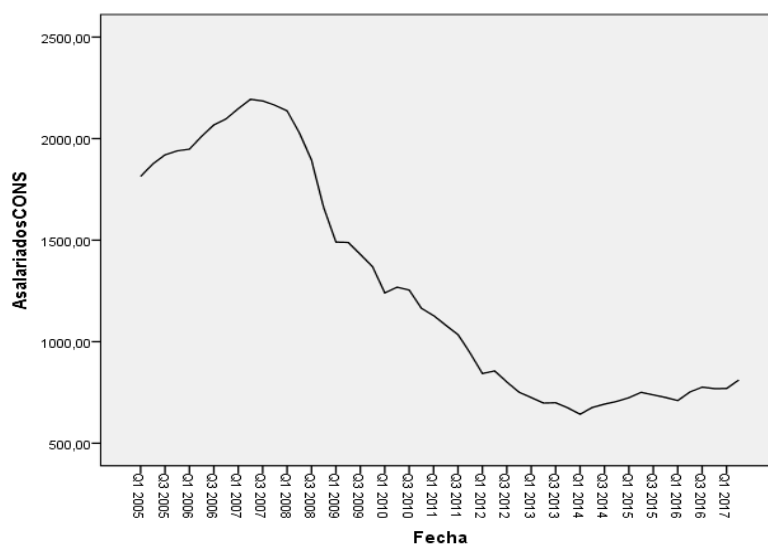


Figura 4.26. Gráfico de secuencia para la serie trimestral asalariados en el sector de la construcción.

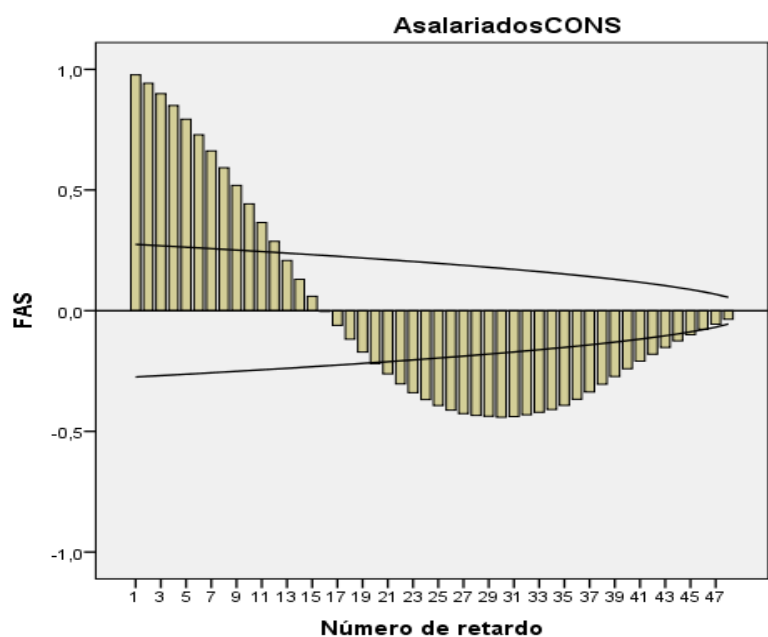


Figura 4.27. FAS para la serie trimestral asalariados en el sector de la construcción.

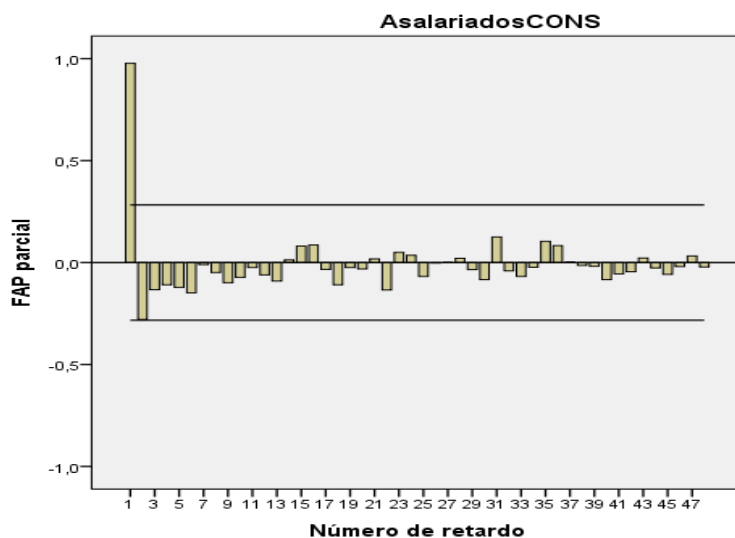


Figura 4.28. FAP de secuencia para la serie trimestral asalariados en el sector de la construcción.

El análisis se realiza de forma análoga a lo visto en las series anteriores. Se propone un modelo ARIMA (1, 1, 0), y los resultados son los siguientes:

$$(1 - 0.6688B)\nabla(z_t) = a_t \quad (4.4)$$

(5.82)

Las FAS y FAP de los residuos no muestran un número de palos que superen las bandas de confianza mayor del 5%. Además, dichos residuos superan los test de Ljung-Box y Box-Pierce, con unos P-valores de 0.5809 y 0.5947 respectivamente. Teniendo en cuenta que los parámetros para el modelo ajustado son significativos, podemos considerar el ajuste como bueno.

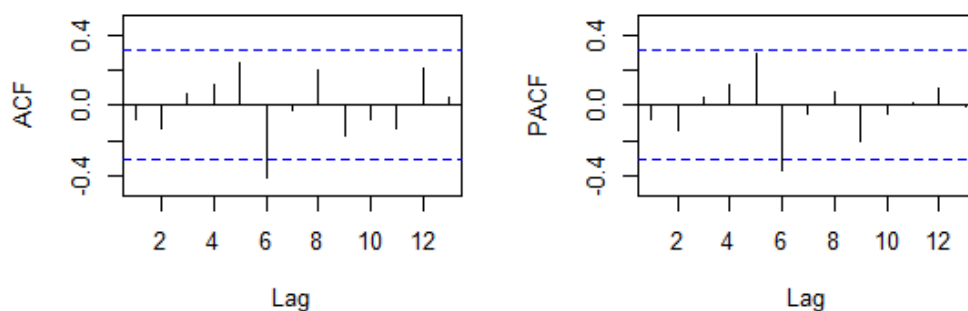


Figura 4.29. FAS y FAP de los residuos del modelo ARIMA ajustado para la serie asalariados en el sector de la construcción.

No se han detectado valores atípicos para esta serie.

4.5 Número de operaciones de compraventa de vivienda nueva.

Esta serie temporal cuenta con datos mensuales desde enero de 2007 hasta agosto de 2017. Las unidades son número de transacciones. En las siguientes imágenes, se puede observar la serie junto a su FAS y FAP.

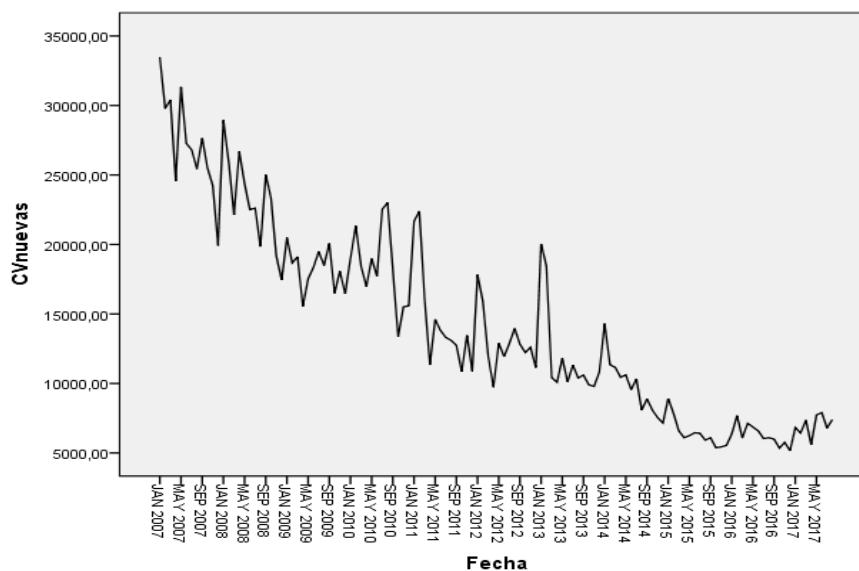


Figura 4.30. Gráfico de secuencia del número de operaciones de compraventa de viviendas nuevas.

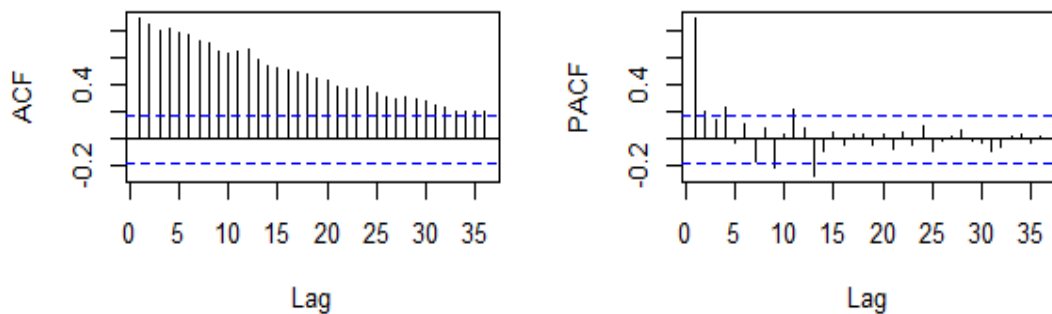


Figura 4.31. FAS y FAP para la serie "número de operaciones de compraventa de viviendas nuevas".

Observando la serie, se decide tomar una diferencia regular y otra estacional para convertir la serie en estacionaria. A raíz de los resultados del test de Dickey-Fuller aumentado (P-valor < 0.01), junto con el aspecto que presenta la serie transformada, la transformación se puede considerar válida. En la imagen inferior se presenta la serie transformada en estacionara, junto a su FAS y FAP.

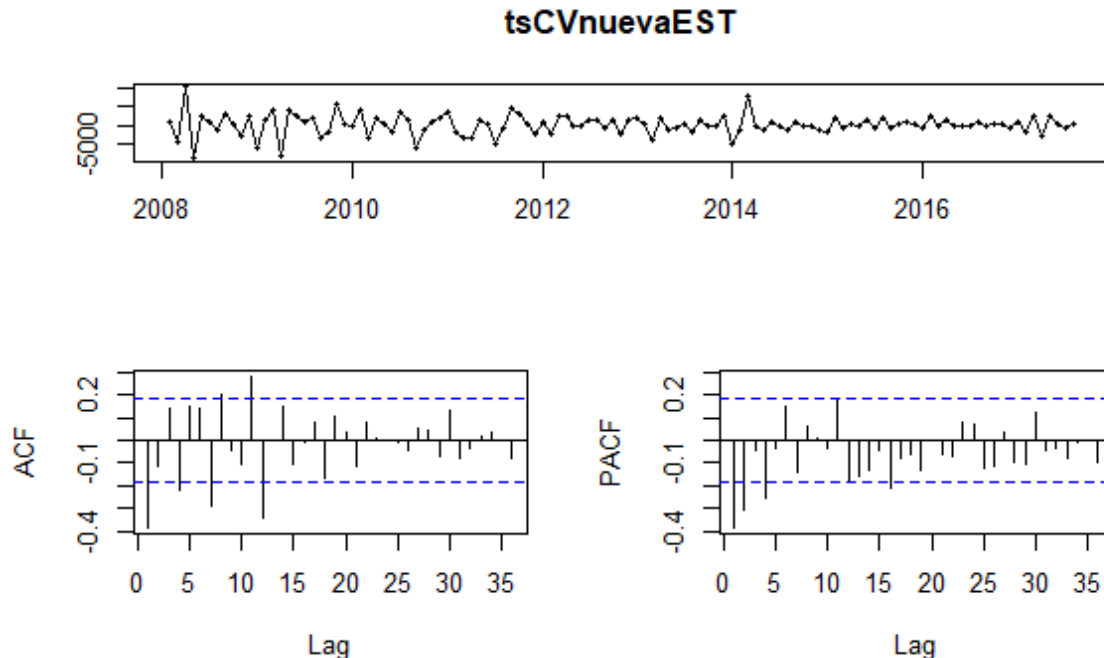


Figura 4.32. Serie de las operaciones de compraventa de vivienda nueva transformada en estacionaria junto a su FAS y FAP.

Se propone ajustar un modelo ARIMA (2,1,0) x (1,1,0). Tras realizar el ajuste, la ecuación del modelo ajustado para la serie es la siguiente:

$$(1 + 0.3078B^{12})(1 + 0.3961B + 0.3151B^2)\nabla\nabla_{12}z_t = a_t \quad (4.5)$$

(-3.31) (-4.35) (-3.48)

Todos los parámetros del modelo resultan significativos. Tal y como muestra la imagen inferior, la FAS y la FAP de los residuos del modelo no presentan palos que superen las bandas de confianza.

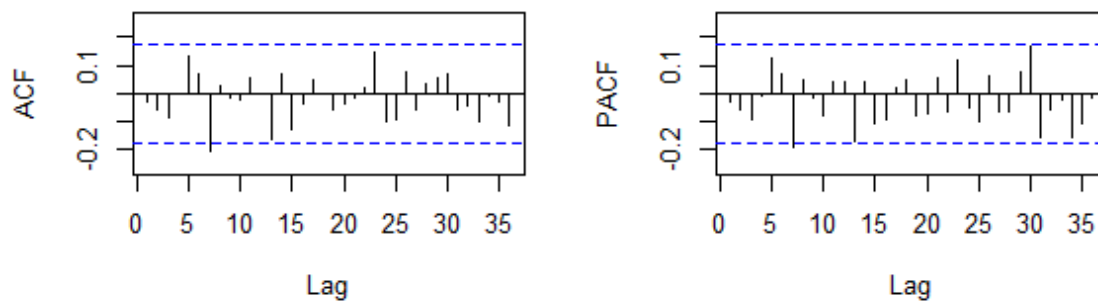


Figura 4.33. FAS y FAP de los residuos generados por el modelo.

Estos residuos además superan los test de Box-Pierce y Ljung-Box, con P-valores de 0.7413 y 0.7442 respectivamente. Podemos concluir que el ajuste es bueno.

A continuación, se presenta una tabla que recoge detalladamente todos los atípicos detectados para esta serie temporal.

Número	Tipo	Posición	Fecha	Magnitud	t
1	AO	16	04/2008	7114	5.305
2	LS	45	09/2010	-5127	-3.645

Tabla 4.7. Atípicos detectados para la serie del número de operaciones compraventa de viviendas nuevas.

El siguiente gráfico muestra los efectos de estos atípicos sobre la serie. .

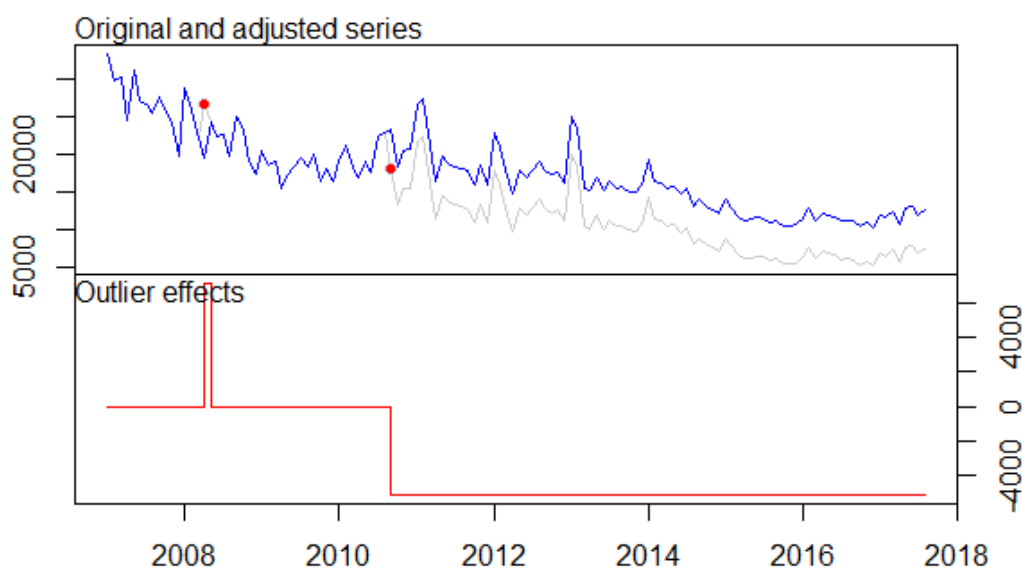


Figura 4.34. Efecto de los atípicos sobre la serie número de operaciones de compraventa de vivienda nueva.

4.6 Número de operaciones de compraventa de vivienda de segunda mano.

La serie analizada muestra datos mensuales que comienzan en enero de 2007 y terminan en agosto de 2017. Su gráfico de secuencia se presenta a continuación, junto a las FAS y FAP.

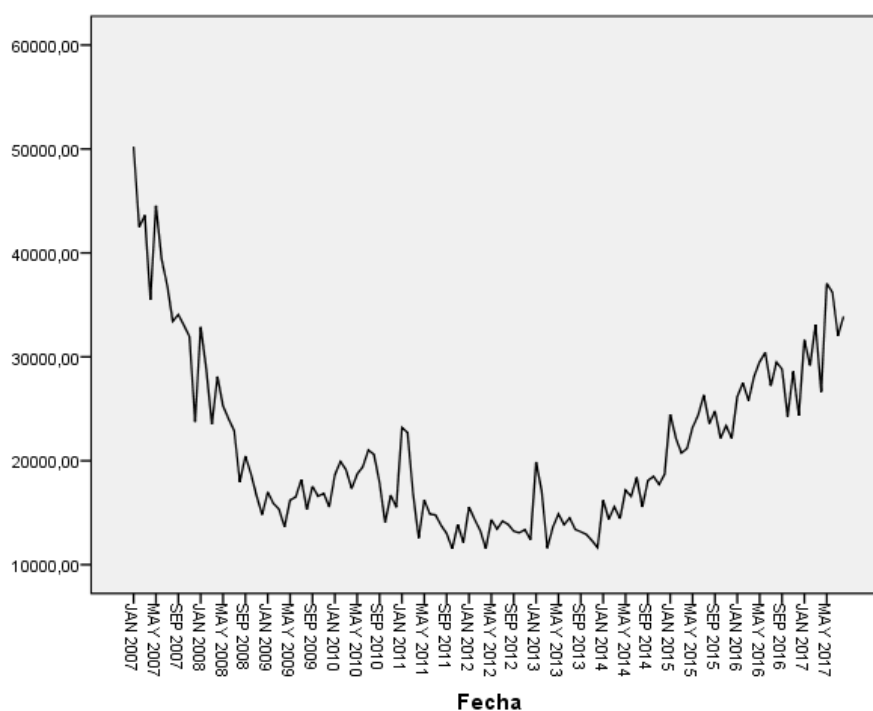


Figura 4.35. Gráfico de secuencia de la serie operaciones de compraventa de viviendas de 2ª mano.

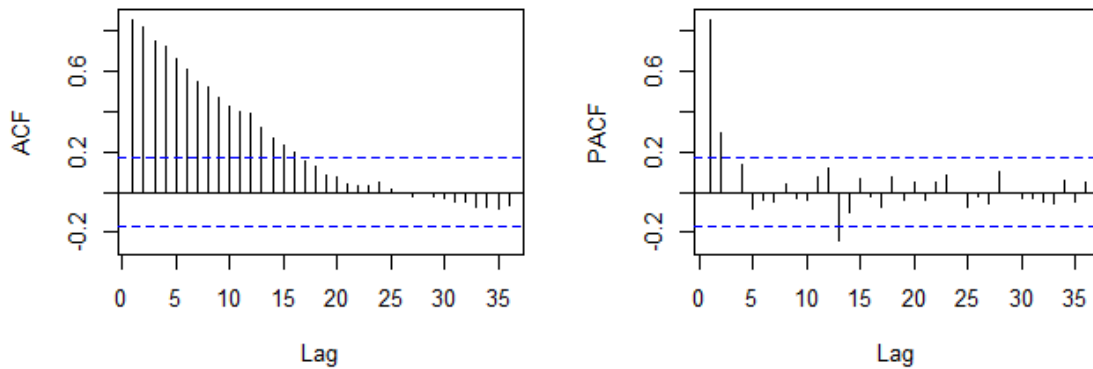


Figura 4.36. FAS y FAP para el número de operaciones de compraventa de viviendas de segunda mano.

Con el objetivo de transformar la serie en estacionaria, se toma una diferencia regular y otra estacional. La serie transformada junto a sus funciones de autocorrelación se muestran en la imagen inferior. El test aumentado de Dickey-Fuller arroja un P-Valor menor que 0.01, por lo que la serie transformada se puede considerar estacionaria.

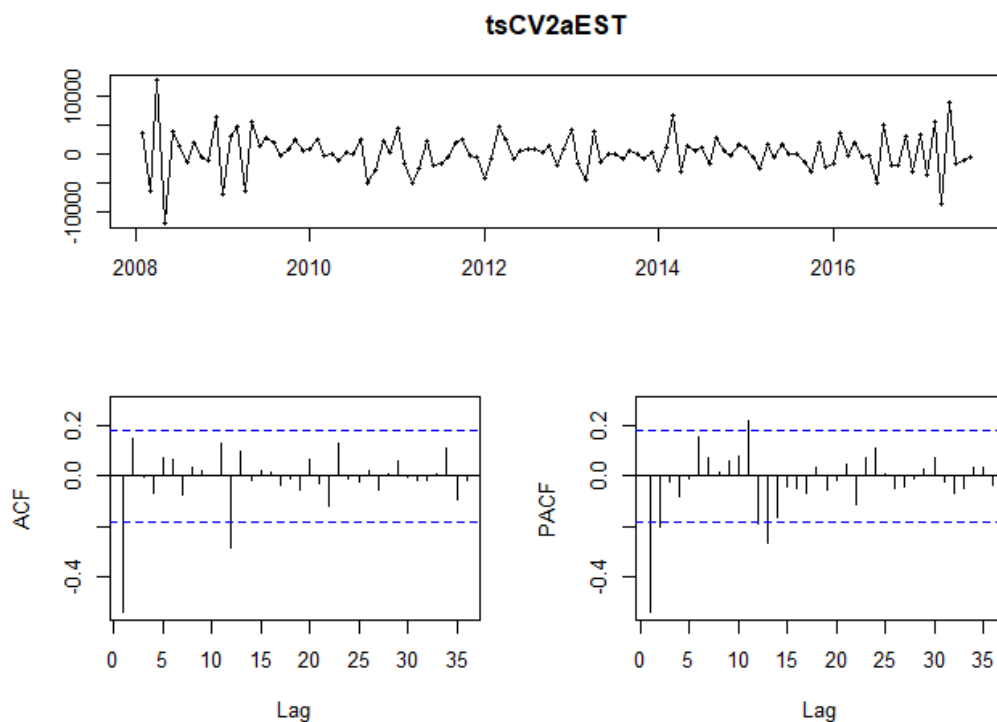


Figura 4.37. Serie del número de operaciones de compraventa de viviendas de segunda mano transformada en estacionaria, junto a su FAS y FAP.

A continuación, se decide ajustar un modelo ARIMA (1,1,1) x (1,1,0). Tras realizar el ajuste, la ecuación para el modelo elegido es la siguiente:

$$(1 + 0.2806B^{12})(1 + 0.3629B)\nabla\nabla_{12}z_t = (1 - 0.2588B)a_t \quad (4.6)$$

(-7.88)
(-4.53)
(6.23)

Todos los parámetros del modelo resultan significativos. Se procede a comprobar la bondad el ajuste estudiando los residuos generados por el modelo. En la imagen inferior vemos como no hay un número significativo de palos que superan las bandas de confianza. Los residuos superan los test de Ljung.Box y Box-pierce con respectivos P-valores de 0.519 y 0.5238. A raíz de estos resultados, se puede considerar el ajuste como bueno.

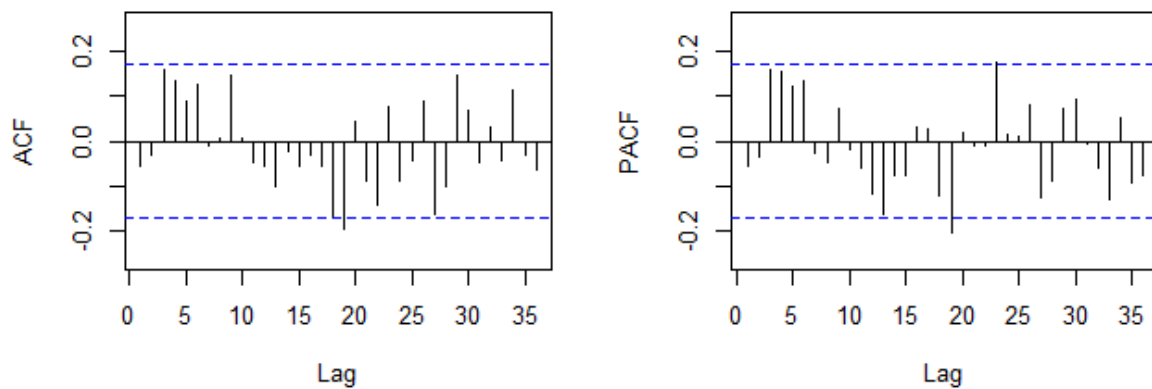


Figura 4.38. FAS y FAP de los residuos generados para el modelo.

La tabla que se presenta bajo este párrafo recoge todos los atípicos detectados para esta serie.

Número	Tipo	Posición	Fecha	Magnitud	t
1	AO	16	04/2008	6438	3.962
2	TC	49	01/2011	6049	3.871

Tabla 4.8. Atípicos para la serie operaciones de compraventa de vivienda de 2ª mano.

A continuación, se presenta un gráfico mostrando el efecto de dichos atípicos sobre la serie.

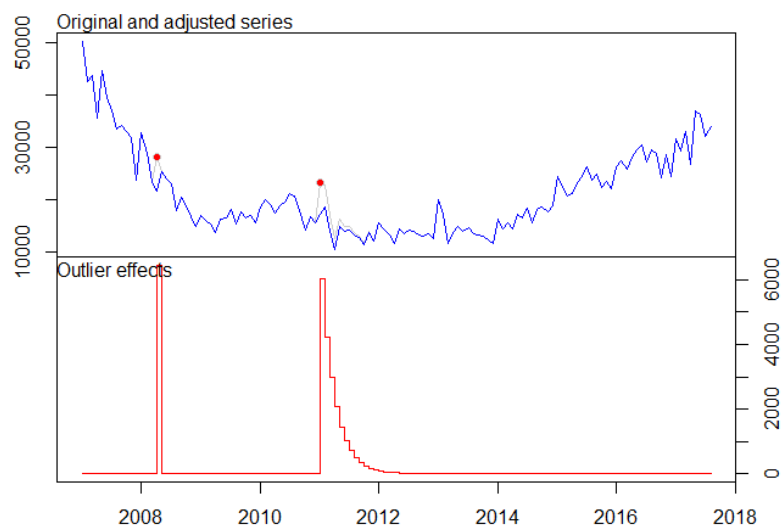


Figura 4.39. Efecto de los outliers sobre la serie del número de operaciones de compraventa de viviendas de segunda mano.

4.7 Hipotecas totales constituidas en España.

Esta serie recoge los datos de la suma total de créditos hipotecarios concedidos sobre todos los tipos de finca. Los datos registrados tienen periodicidad mensual, comienzan en enero de 2003 y concluyen en septiembre de 2017. El gráfico de secuencia de esta serie temporal se muestra a continuación, junto a su FAS y FAP.

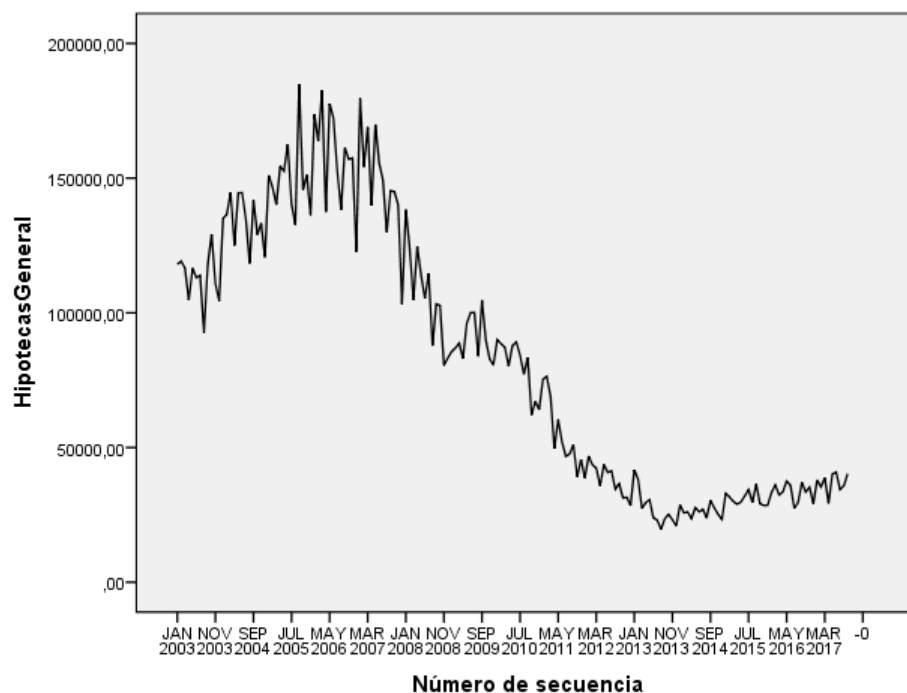


Figura 4.40. Gráfico de secuencia de la serie temporal hipotecas totales constituidas en España.

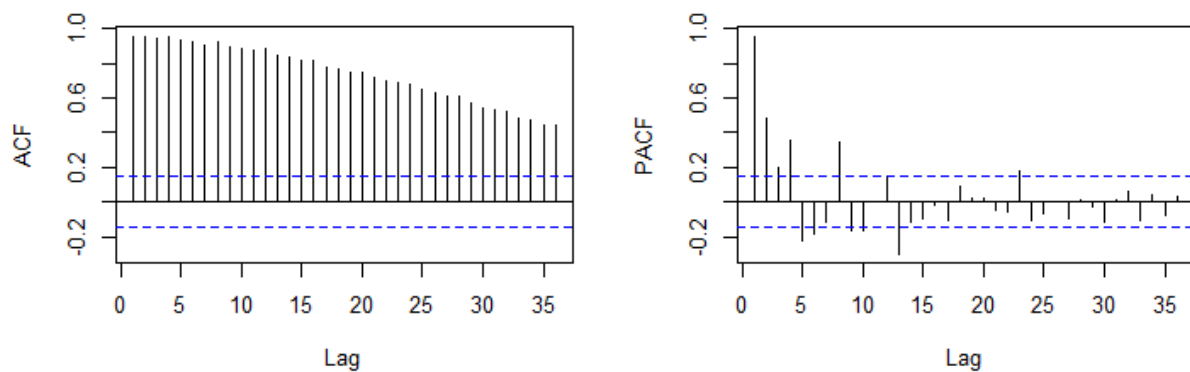


Figura 4.41. Funciones de autocorrelación para la serie de hipotecas constituidas en España.

Se toma una diferencia regular y otra estacional para transformar la serie en estacionaria. La serie transformada en estacionaria, junto a su FAS y FAP, se muestran en la imagen inferior.

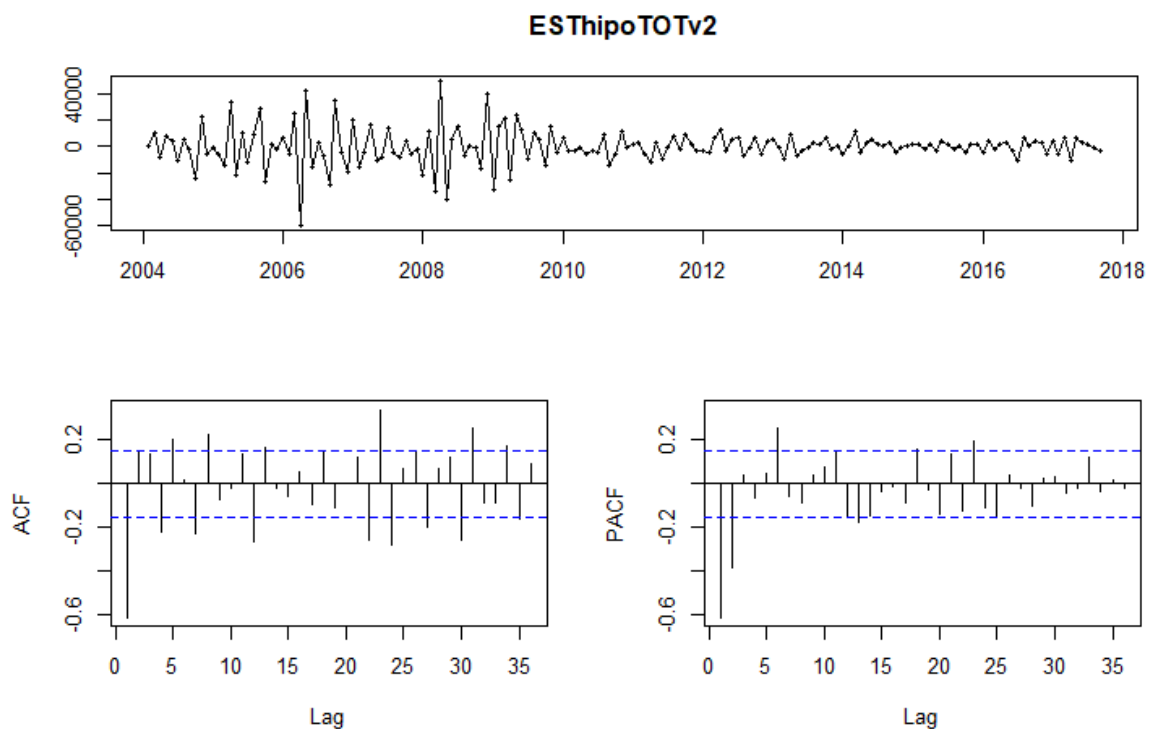


Figura 4.42. Serie hipotecas totales constituidas en España transformada en estacionaria junto a sus funciones de autocorrelación FAS y FAP.

Estudiando la estructura de las funciones de la FAS y la FAP, se propone ajustar un modelo ARIMA (1, 1, 1) x (1, 1, 1). Tras realizar el ajuste, la expresión que regula el comportamiento de la serie en base a este modelo es la siguiente

$$(1 - 0.420B^{12})(1 + 0.2976B)\nabla\nabla_{12}z_t = (1 - 0.703B^{12})(1 - 0.491B)a_t \quad (4.7)$$

(2.584) (-3.173) (5.337) (5.730)

Estudiamos las FAS y FAP de los residuos para determinar si pueden ser considerados como ruido blanco. Como no hay un número significativo de palos que supere las bandas de confianza, y los residuos superan los test de Ljung-Box y Box-Pierce (P-valores 0.7982 y 0.8013), podemos considerar el ajuste como bueno. Además, todos los parámetros del modelo son significativos. A continuación, se muestra la FAS y FAP de los residuos.

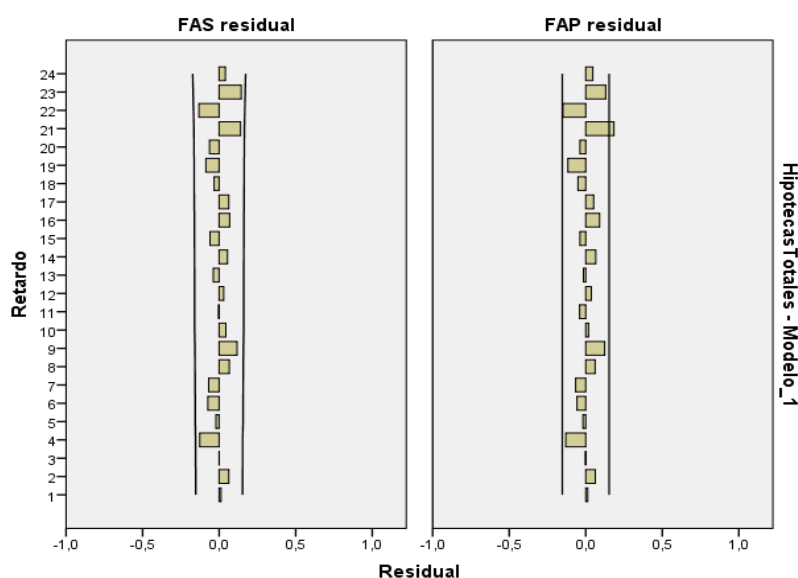


Figura 4.43. FAS y FAP de los residuos del modelo ARIMA para hipotecas totales constituidas en España.

En la tabla inferior, se muestra una tabla con todos los outliers detectados para el programa, indicando su tipo, localización en el tiempo y magnitud:

Número	Tipo	Fecha	Posición	Magnitud	Sig. (t)
1	Aditivo	09/2005	33	30602	7.200
2	C.T.	03/2006	39	20293	4.085
3	Aditivo	04/2006	40	-35489	-7.315
4	C. de Nivel	03/2008	63	-28473	-5.701
5	Aditivo	04/2008	64	35979	7.492
6	SLS*	12/2008	72	20776	6.302
7	SLS*	08/2010	92	17175	5.550

Tabla 4.9. Atípicos detectados en R para la serie temporal total de hipotecas constituidas en España.

Más abajo, podemos ver el gráfico TSO para esta serie, en la que se observa el efecto de los atípicos sobre la serie:

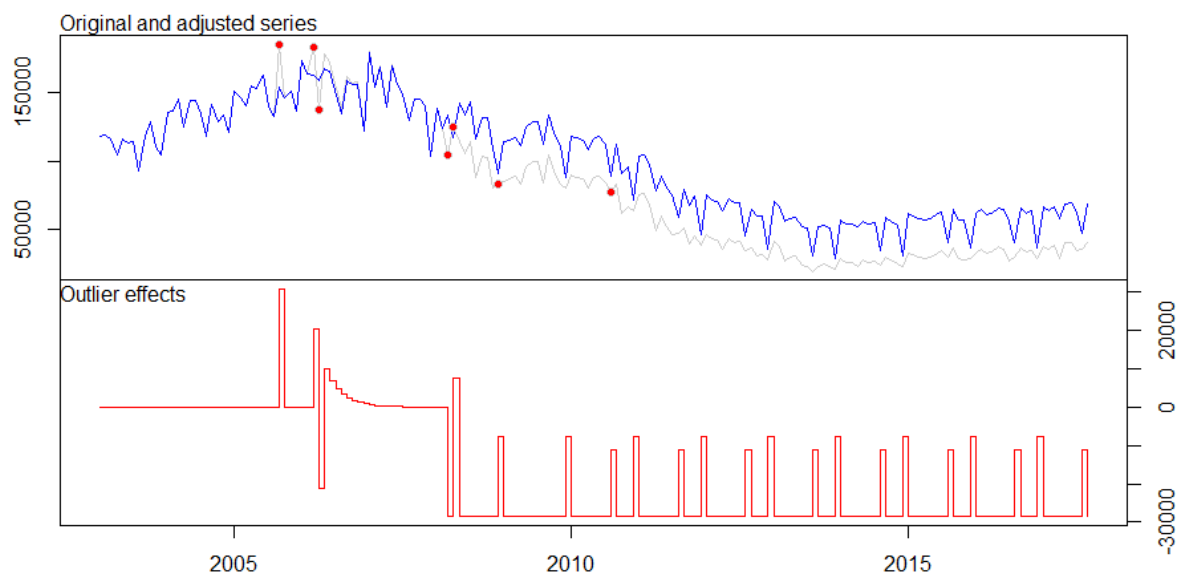


Figura 4.44. Efectos de los outliers sobre el total de las hipotecas constituidas en España.

4.8 Índice de producción industrial (IPI) de la industria metalúrgica en España.

Los IPI, siglas de “índice de producción industrial”, son indicadores económicos de coyuntura de la actividad productiva de algún sector industrial. En este caso, vamos a analizar el IPI para la industria metalúrgica. Esta serie tiene una periodicidad mensual, y los datos registrados van desde enero de 1995 hasta agosto de 2017. A continuación, se muestra la serie temporal junto a sus FAS y FAP.

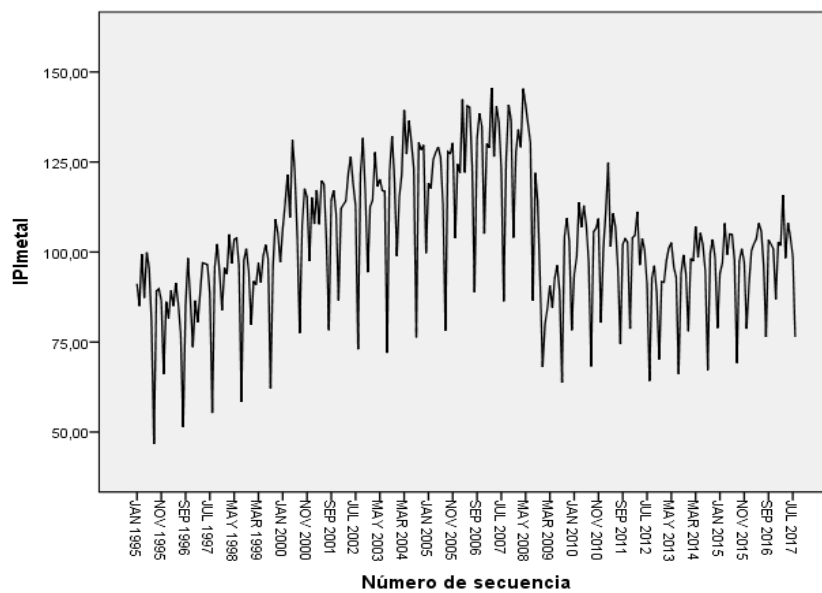


Figura 4.45. Gráfico de secuencia del IPI en la industria metalúrgica.

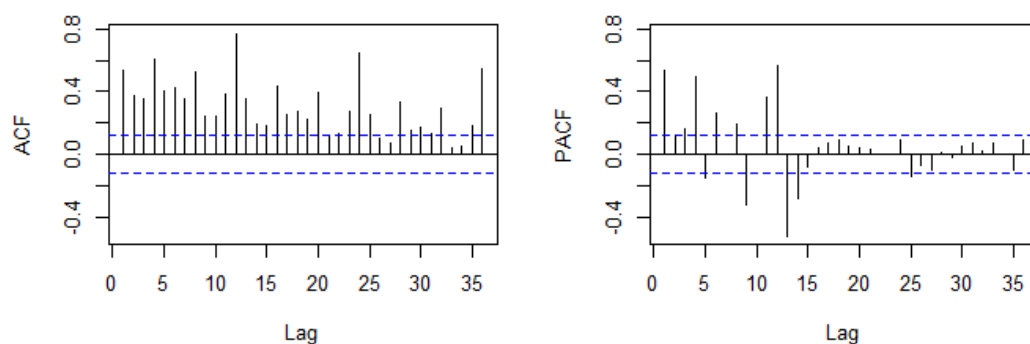


Figura 4.46. FAS y FAP de la serie del IPI en la industria metalúrgica.

Para transformar la serie en estacionaria, se toma una diferencia regular y otra estacional. Tras esta transformación, el test aumentado de Dickey-Fuller arroja un P-valor menor que

0.01, por lo que la serie se puede considerar estacionaria. Las gráficas inferiores muestran la serie ya transformada junto a sus FAS y FAP.

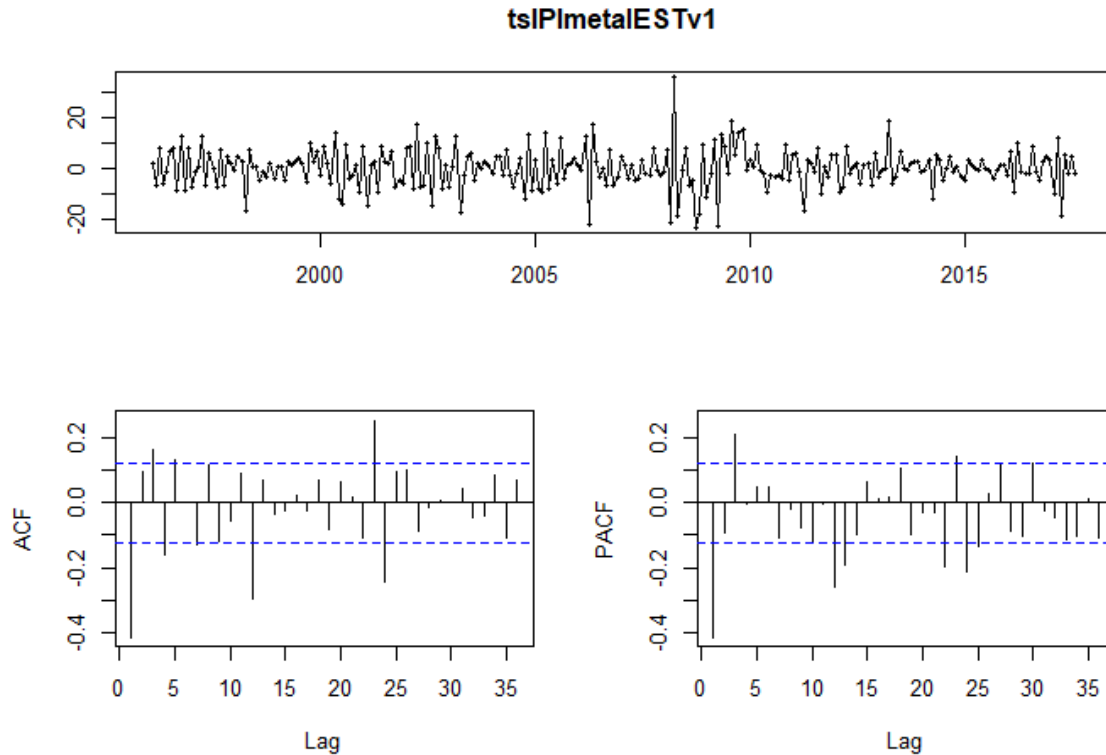


Figura 4.47. Serie IPI de la industria metalúrgica transformada en estacionaria.

Se propone ajustar un modelo ARIMA (2, 1, 0) x (0, 1, 1). Tras realizar el ajuste, la ecuación que explica el comportamiento de la serie analizada es la siguiente:

$$(1 + 0.6325B + 0.2624B^2)\nabla\nabla_{12}z_t = (1 - 0.8482B^{12})a_t \quad (4.8)$$

(-10.07) (-4.308) (15.88)

Todos los parámetros del modelo son significativos. En la imagen inferior se puede observar que no hay un número significativo de palos que superen las bandas de confianza en las FAS y FAP de los residuos. Además, dichos residuos superan los test de Ljung-Box y Box-Pierce con sendos P-valores de 0.9308 y 0.9305. Podemos concluir, por lo tanto, que el ajuste es bueno.

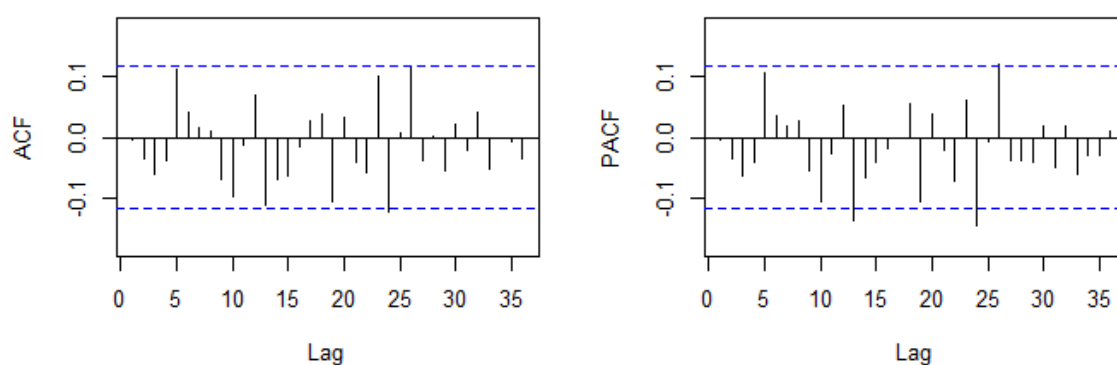


Figura 4.48. Residuos del modelo ajustado para IPI en la industria metalúrgica.

A continuación, se muestra una tabla con todos los outliers detectados para esta serie:

Número	Tipo	Posición	Fecha	Magnitud	"t"
1	SLS	84	12/2001	-10.8589	-4.322
2	TC	160	04/2008	18.0527	4.434
3	SLS	164	07/2008	12.6406	5.340
4	LS	167	11/2008	-23,8899	-6.080

Tabla 4.10. Atípicos detectados en R para la serie IPI en la industria metalúrgica.

La siguiente imagen muestra los efectos de estos atípicos sobre la serie:

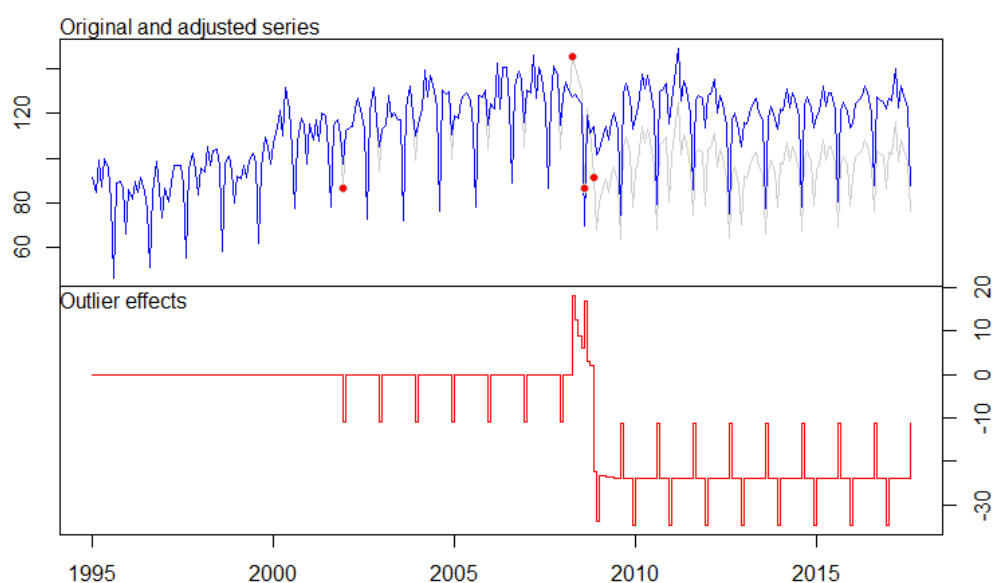


Figura 4.49. Efectos de los atípicos sobre la serie IPI en la industria metalúrgica.

4.9 Producción interna de acero en España.

La producción interna de acero mide la cantidad de toneladas métricas de acero producidas en España a nivel nacional, declaradas por los grandes fabricantes. La serie presenta datos mensuales, cuyo registro comienza en enero de 1995 y finaliza en mayo de 2017. Las unidades son toneladas métricas de acero producido. El gráfico de secuencia de la serie se muestra a continuación, así como sus FAS y FAP.

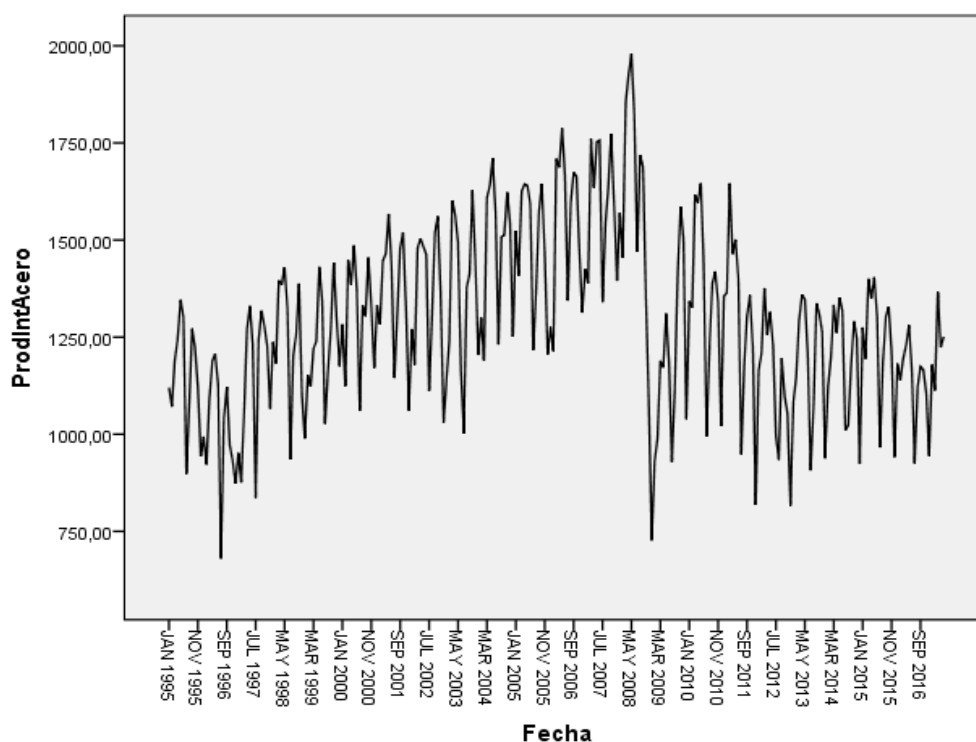


Figura 4.50. Gráfico de secuencia de la serie temporal producción interna de acero.

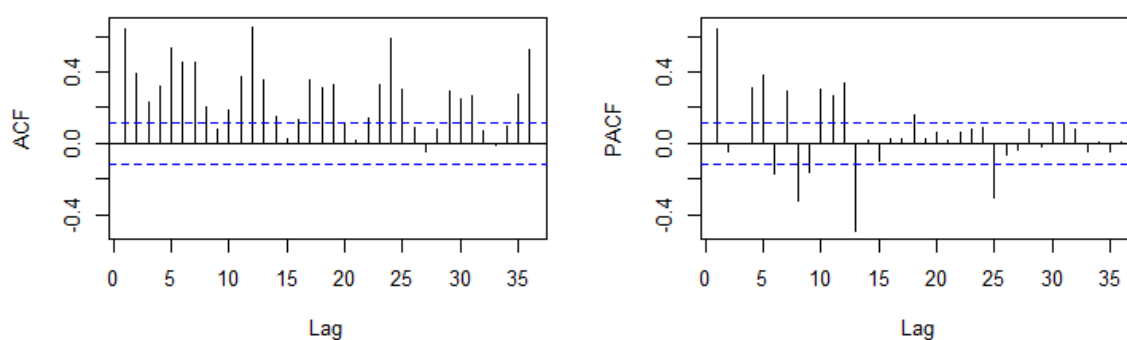


Figura 4.51. FAS y FAP junto al gráfico de la serie producción interna de acero.

Para transformar la serie en estacionaria, se toma una diferencia regular y otra estacional, y se somete a la serie transformada al test aumentado de Dickey-Fuller. El resultado del test es un P-Valor menor que 0.01, por lo que la serie transformada puede considerarse estacionaria. Las gráficas a continuación de este párrafo muestran la serie transformada, junto a sus FAS y FAP.

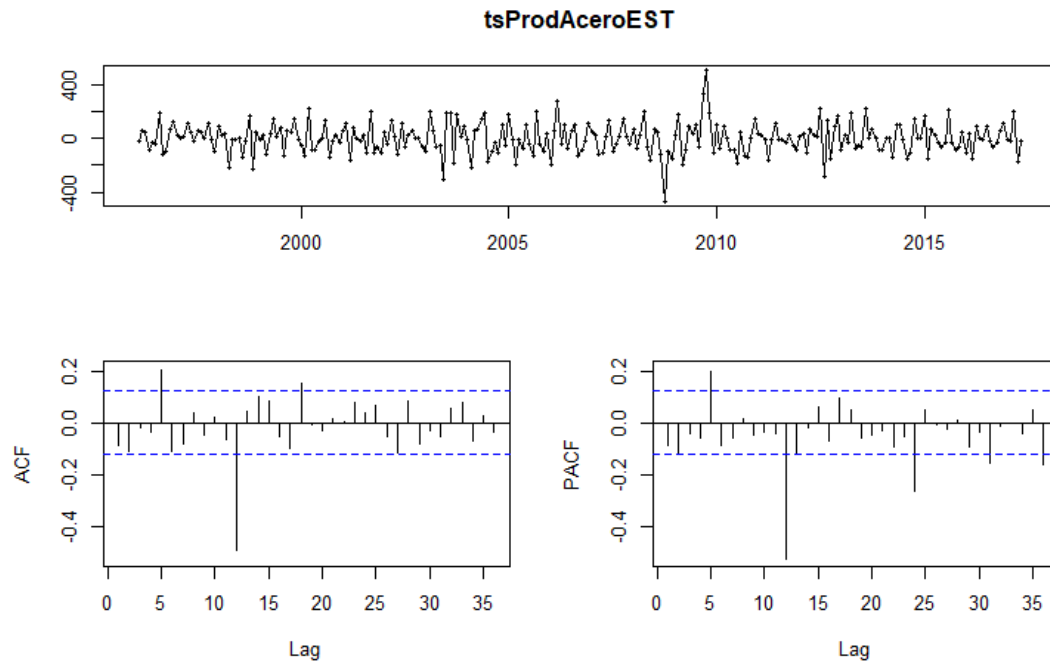


Figura 4.52. Serie producción interna de acero tras tomar una diferencia regular y otra estacional, junto a su FAS y FAP.

Se propone ajustar un modelo ARIMA (0, 1, 2) x (2, 1, 0). Tras llevar a cabo el ajuste, la ecuación que explica el comportamiento de la serie en base al modelo elegido es la siguiente:

$$(1 + 0.5877B^{12} + 0.2519B^{24})\nabla\nabla_{12}z_t = (1 - 0.4556B - 0.15B^2)a_t \quad (4.9)$$

(-8.97) (-4.03)
(6.95) (2.29)

Todos los parámetros del modelo resultan significativos. Los residuos generados por el modelo no presentan estructura ni palos significativos en las funciones de correlación, tal y como muestra la imagen inferior. Además, pasan holgadamente los test de Ljung-Box y Box-Pierce con respectivos P-valores de 0.9386 y 0.9389. A raíz de los resultados obtenidos, el ajuste del modelo seleccionado puede considerarse como bueno.

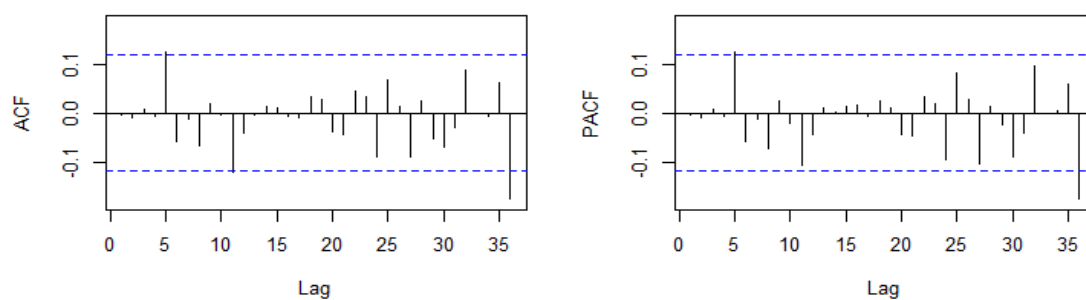


Figura 4.53. Residuos generados para el ajuste del modelo ARIMA propuesto para la serie producción interna de acero.

A continuación, se muestra una tabla con todos los atípicos detectados para esta serie:

Número	Tipo	Posición	Fecha	Magnitud	"t"
1	TC	102	07/2003	-268.0	-4.200
2	LS	166	10/2008	-545.9	-9.195
3	LS	177	09/2009	282.0	4.856

Tabla 4.11. Atípicos detectados para la serie producción interna de acero.

En la imagen inferior se muestra un gráfico que recoge el impacto de estos outliers en el comportamiento de la serie:

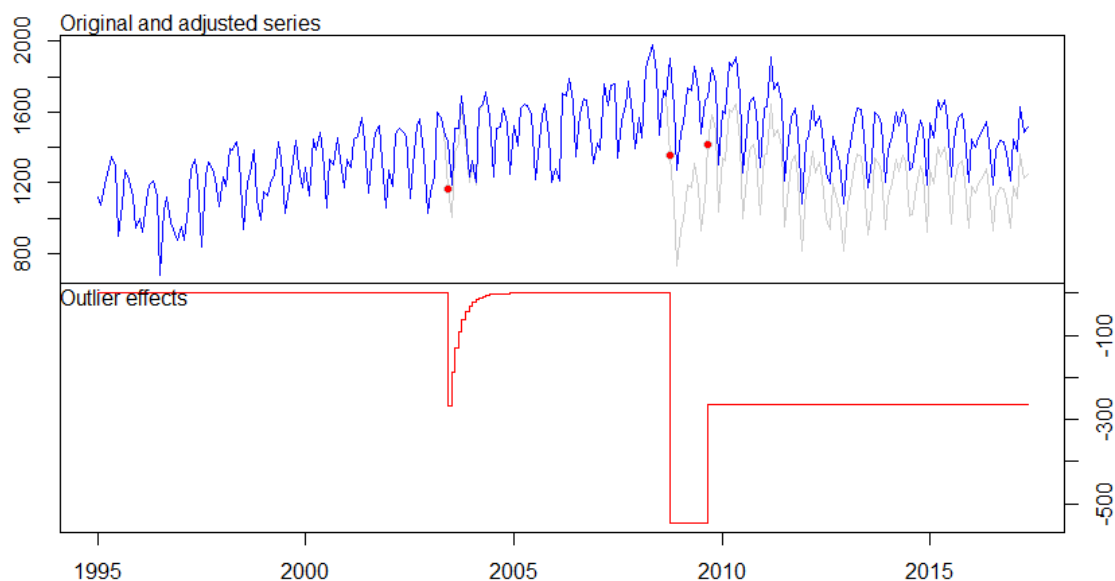


Figura 4.54. Efecto de los atípicos en la serie producción interna de acero.

4.10 Exportaciones de acero.

Esta serie temporal recoge los datos sobre las toneladas métricas de acero que son exportadas desde España hacia cualquier país del extranjero. Esta serie temporal muestra datos comprendidos entre enero de 1995 y abril de 2017, ambos inclusive. A continuación, se presenta el gráfico de secuencia de la serie, junto a los gráficos de sus FAS y FAP.

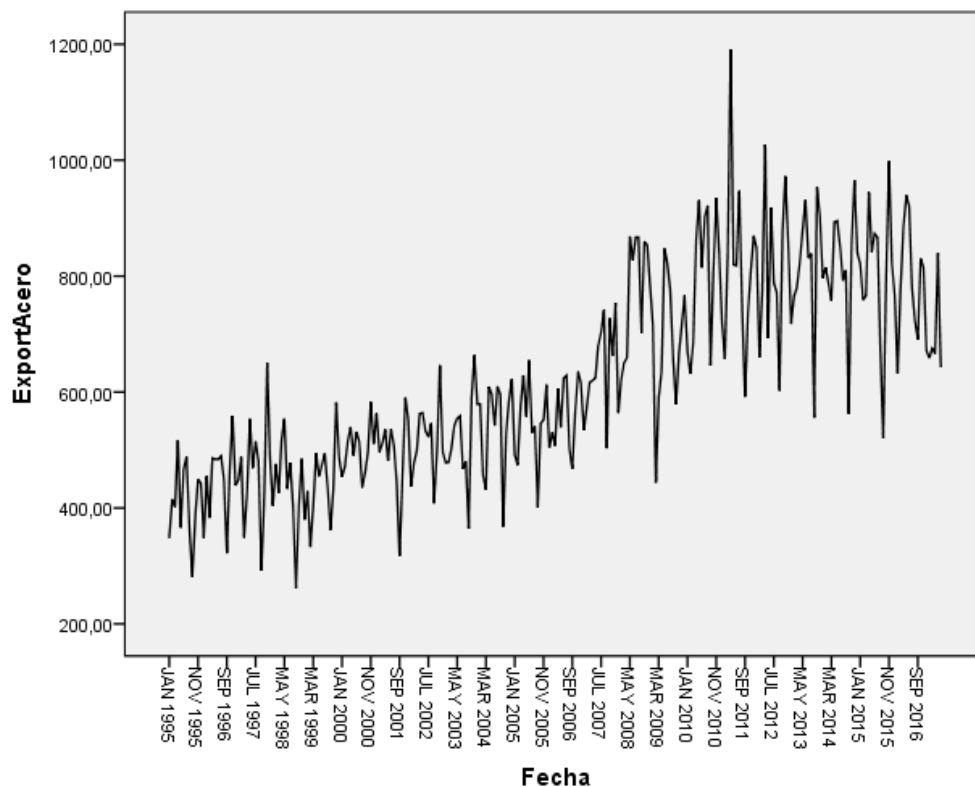


Figura 4.55. Gráfico de secuencia de la serie exportación de acero.

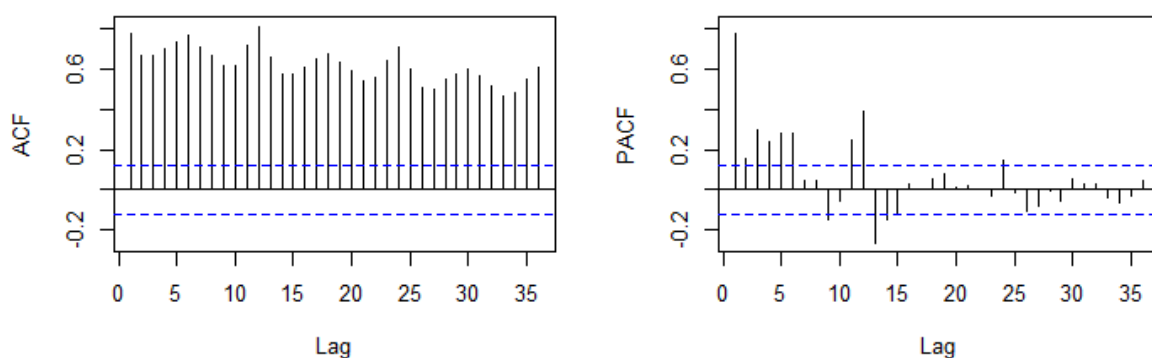


Figura 4.56. Funciones de autocorrelación para la serie exportaciones de acero.

Observando la serie, se determina que es necesario tomar una diferencia regular y otra estacional para transformar la serie en estacionaria. Tras esta transformación, el resultado del test aumentado de Dickey-Fuller arroja un P-valor menor que 0.01, por lo que la serie se puede considerar estacionaria. A continuación, se muestra la serie transformada junto a sus FAS y FAP.

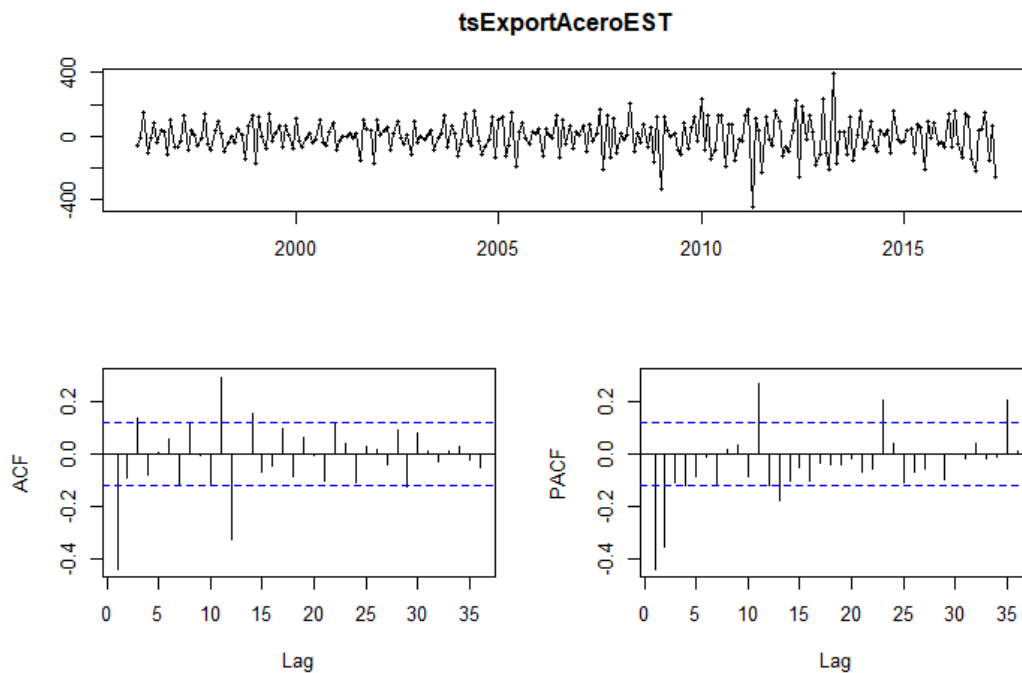


Figura 4.57. Serie exportaciones de acero transformada en estacionara junto a su FAS y FAP.

Una vez transformada en estacionaria, podemos ajustar un modelo ARIMA sobre la serie. Estudiando la estructura de la FAS y la FAP, se propone un modelo ARIMA (0, 1, 1) x (0, 1, 1). Tras llevar a cabo el ajuste, la ecuación que explica el comportamiento de la serie en base al modelo elegido es la siguiente:

$$\nabla \nabla_{12} z_t = (1 - 0.8145B^{12})(1 - 0.7739B)a_t \quad (4.10)$$

$$(14.97) \quad (17.71)$$

Todos los parámetros del modelo resultan significativos. Una vez realizado el ajuste, se estudian los residuos generados. Como se aprecia en la imagen de la FAS y la FAP de los

residuos, no hay un número significativo de palos que superen los límites de control. Además, los test de Ljung-Box y Box-Pierce son superados con éxito (P-valores de 0.6965 y 0.6981 respectivamente).

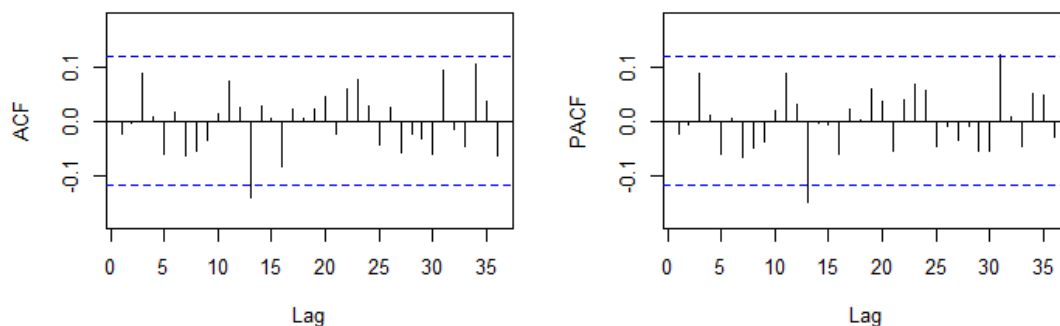


Figura 4.58. Residuos generados en el ajuste para la serie exportaciones de acero junto a su FAS y FAP.

En la tabla inferior se encuentran expuestos detalladamente todos los atípicos que han sido detectados durante el ajuste:

Número	Tipo	Posición	Fecha	Magnitud	"t"
1	TC	169	01/2009	-241.7	-4.739
2	AO	195	03/2011	324.2	5.582

Tabla 4.12. Atípicos detectados para la serie exportaciones de acero.

El gráfico adjunto muestra los efectos de los outliers sobre la serie:

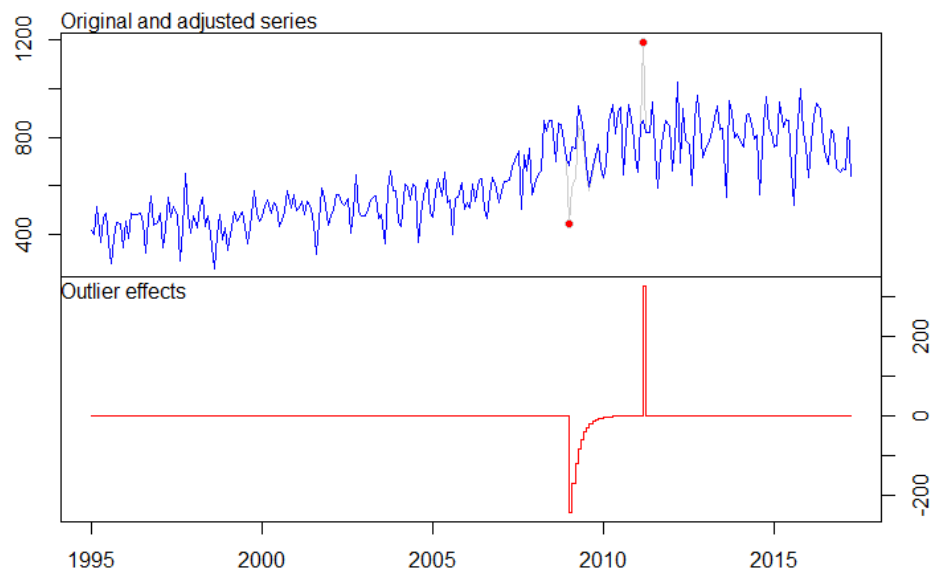


Figura 4.59. Efecto de los atípicos sobre la serie exportaciones de acero.

4.11 Importaciones de acero en España.

Esta serie temporal recoge los datos sobre las toneladas métricas de acero que son importadas en España desde cualquier país del extranjero. La serie contiene datos mensuales desde enero de 1995 hasta abril de 2017. Las imágenes inferiores muestran el gráfico de secuencia de la serie junto a su FAS y FAP

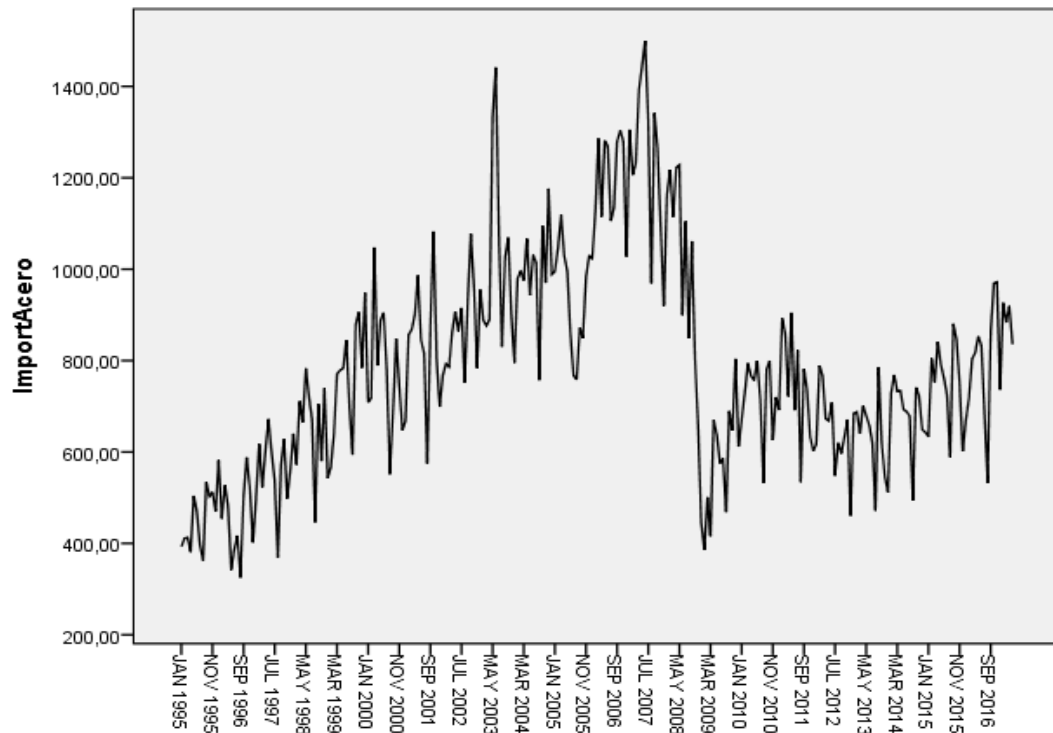


Figura 4.60. Gráfico de secuencia para la serie importaciones de acero.

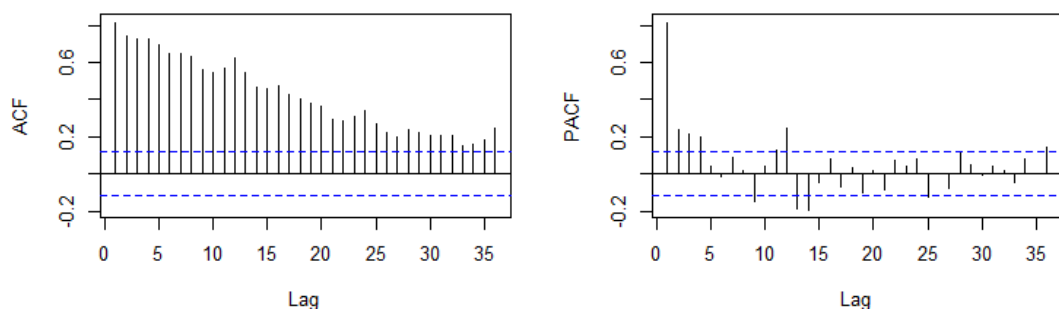


Figura 4.61. FAS y FAP de la serie importaciones de acero junto a su gráfico de secuencia.

Se toma una diferencia regular y otra estacional para transformar la serie en estacionaria. A continuación, se muestra la serie ya transformado junto a sus funciones de autocorrelación simple y parcial.

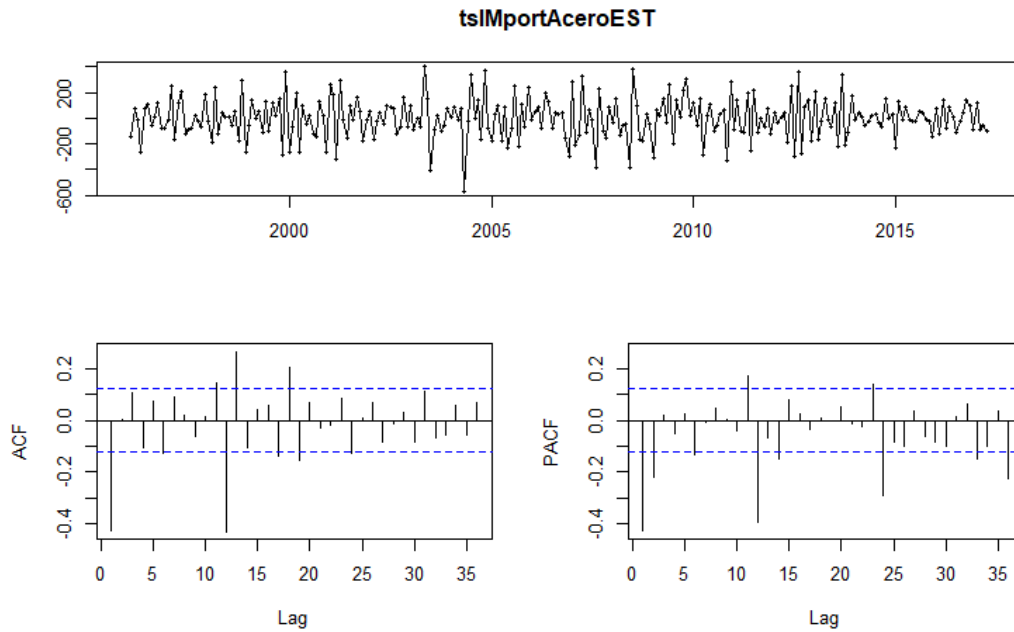


Figura 4.62. Serie importaciones de acero transformada en estacionaria.

A raíz del resultado del test de Dickey-Fuller (P-Valor menor que 0.01), la serie transformada puede considerarse estacionaria. Se decide ajustar un modelo ARIMA (0, 1, 1) x (2, 1, 0). Tras realizar el ajuste, la ecuación que explica el comportamiento de la serie según dicho modelo es la siguiente:

$$(1 + 0.5683B^{12} + 0.4860B^{24})\nabla\nabla_{12}z_t = (1 - 0.6870B)a_t \quad (4.11)$$

(-10.22)

(-8.74)

(14.02)

Todos los parámetros del modelo ajustado resultan significativos. Se comprueba la bondad la bondad del ajuste estudiando los residuos generados por el modelo. En la imagen de la página siguiente, se observa que la FAS y FAP de dichos residuos no muestra un número significativo de palos que superen las bandas de confianza. Además, los test de Ljung-Box y Box-Pierce son superados con éxito, con P-valores de 0.8548 y 0.8556 respectivamente. Teniendo todo esto cuenta, se puede concluir que el ajuste es bueno.

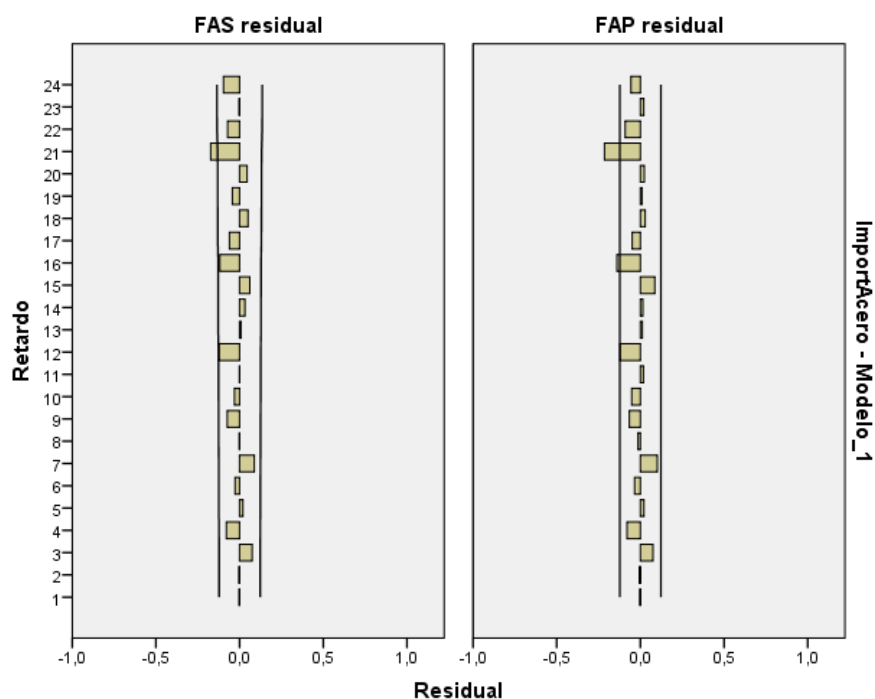


Figura 4.63. FAS y FAP de los residuos del modelo ajustado para la serie importaciones de acero.

La siguiente tabla recoge todos los outliers detectados para el modelo ARIMA ajustado.

Número	Tipo	Posición	Fecha	Magnitud	"t"
1	AO	60	12 1999	316.1	4.426
2	AO	63	03 2000	275.4	3.835
3	TC	101	05 2003	411.9	6.182
4	TC	119	11 2004	311.6	4.642
5	LS	132	12 2005	298.7	5.123
6	LS	167	11 2008	-391.3	-6.499

Tabla 4.13. Atípicos detectados para la serie importaciones de acero.

Además, el gráfico adjunto muestra el efecto de estos outliers en el comportamiento de la serie.

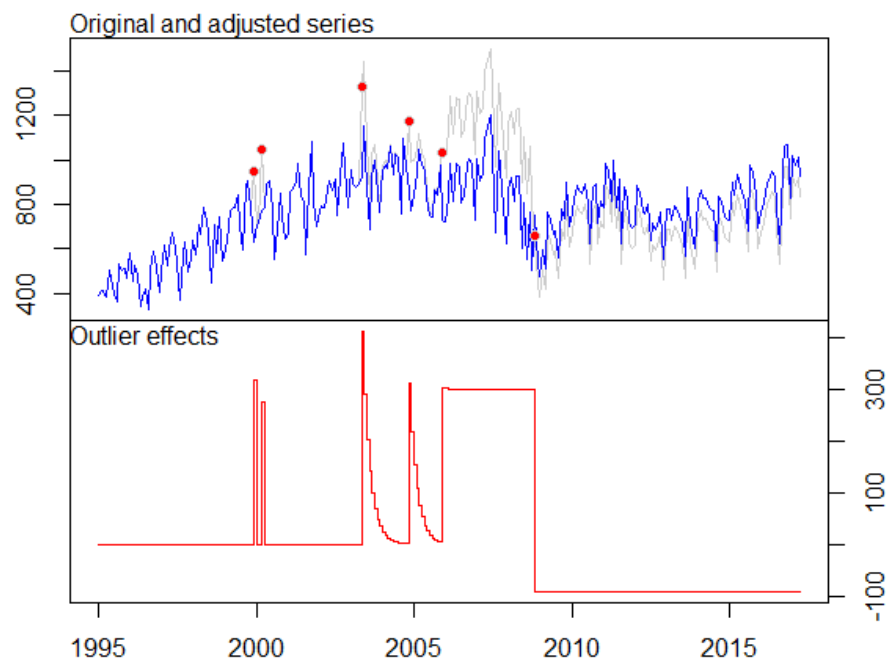


Figura 4.64. Efecto de los atípicos sobre las series de las importaciones de acero.

4.12 Paro total registrado en España.

La serie contiene datos mensuales desde enero de 1996 hasta octubre de 2016. Su gráfico de secuencia se muestra a continuación, acompañado de la FAS y la FAP.

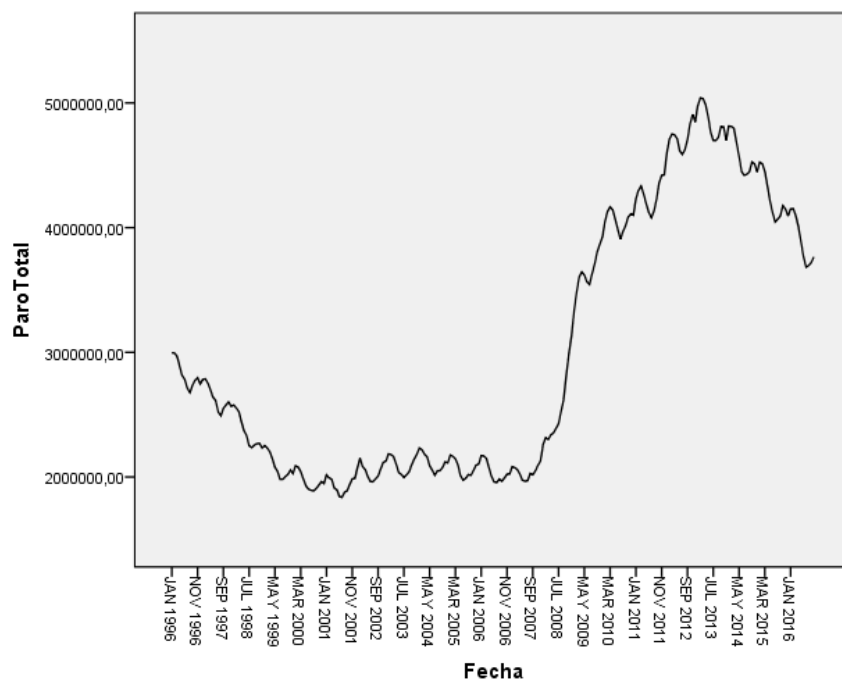


Figura 4.65. Gráfico de secuencia de la serie temporal del paro total registrado.

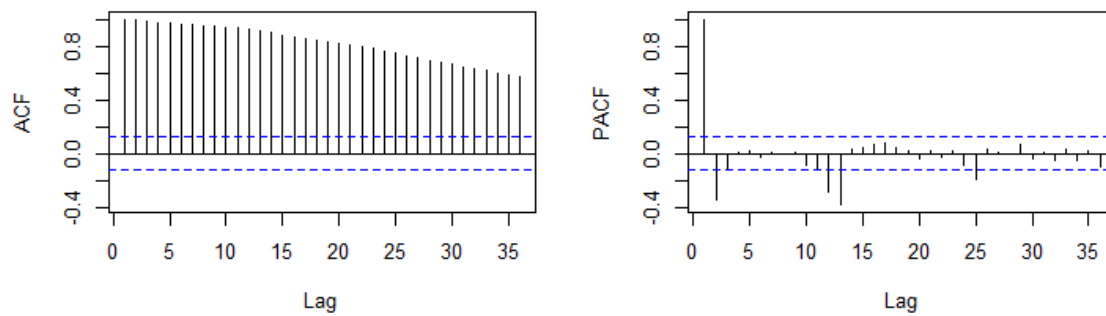


Figura 4.66. FAS y FAP de la serie del paro total registrado, junto a su gráfico de secuencia.

En este caso, la transformación en serie estacionaria requiere de dos diferencias regulares en lugar una sola, además de una diferencia estacional. Para estas transformaciones, el test de Dickey-Fuller corrobora que la serie transformada es estacionaria. A continuación, se muestra la serie transformada junto a su FAS y su FAP.

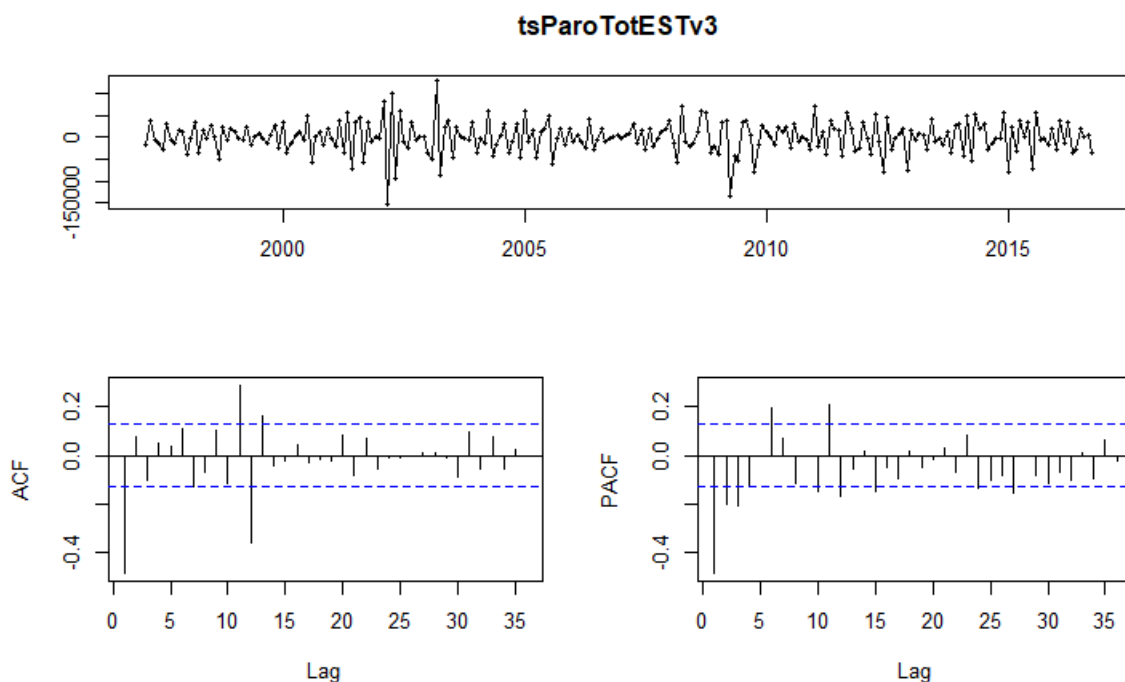


Figura 4.67. Serie del paro total registrado transformada en estacionaria, junto a su FAS y FAP.

Se propone ajustar un modelo ARIMA (0,1,1) x (1,1,1). Tras realizar el ajuste, la ecuación que explica el comportamiento de la serie es la siguiente:

$$(1 - 0.4655B^{12})\nabla^2\nabla_{12}z_t = (1 - 0.7636B^{12})(1 - 0.5593B)a_t \quad (4.12)$$

(2.418)

(4.83)

(9.93)

Tras realizar el ajuste, se estudian los residuos generados por el modelo. Tanto la FAS como la FAP de estos residuos no muestran palos que excedan las bandas de confianza, tal y como muestra la imagen inferior. Además, los residuos superan los test de Ljung-Box y Box-Pierce, con P-valores de 0.7145 y 0.7161. Todos estos hechos nos llevan a considerar el ajuste como bueno.

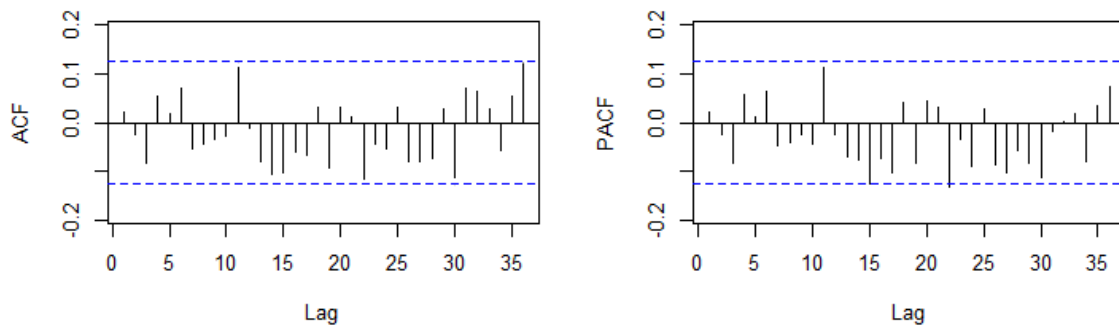


Figura 4.68. Residuos generados por el modelo para la serie del paro total registrado.

A continuación, se muestra una tabla que recoge con detalle todos los atípicos detectados:

Número	Tipo	Posición	Fecha	Magnitud	t
1	AO	74	02/2002	69391	5.888
2	TC	147	03/2008	-11784	-4.619

Tabla 4.14. Atípicos detectados para la serie del paro total.

Por último, se adjunta un gráfico donde se muestra el efecto de estos atípicos sobre la serie:

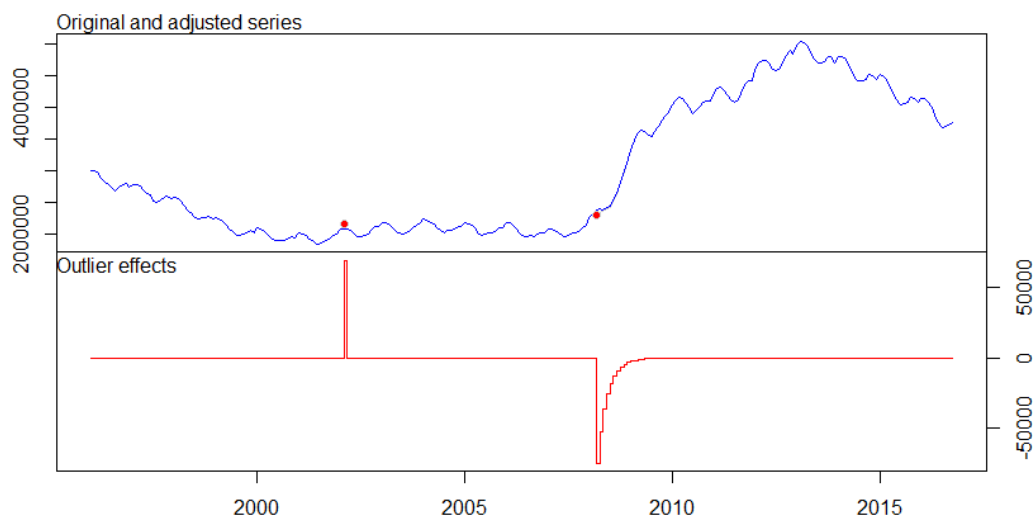


Figura 4.69. Efecto de los atípicos sobre la serie del paro total registrado.

4.13 IPC: índice de precios de consumo.

El Índice de precios de consumo (IPC) es una medida estadística de la evolución de los precios de los bienes y servicios que consume la población residente en viviendas familiares en España. Es un indicador macroeconómico general que aporta información directa de como la situación económica es percibida, asimilada y respondida por la población. La serie muestra datos mensuales comenzando en enero de 2002 y finalizando en agosto de 2017. A continuación, se muestra el gráfico de secuencia de la serie, acompañado de su FAS y FAP.

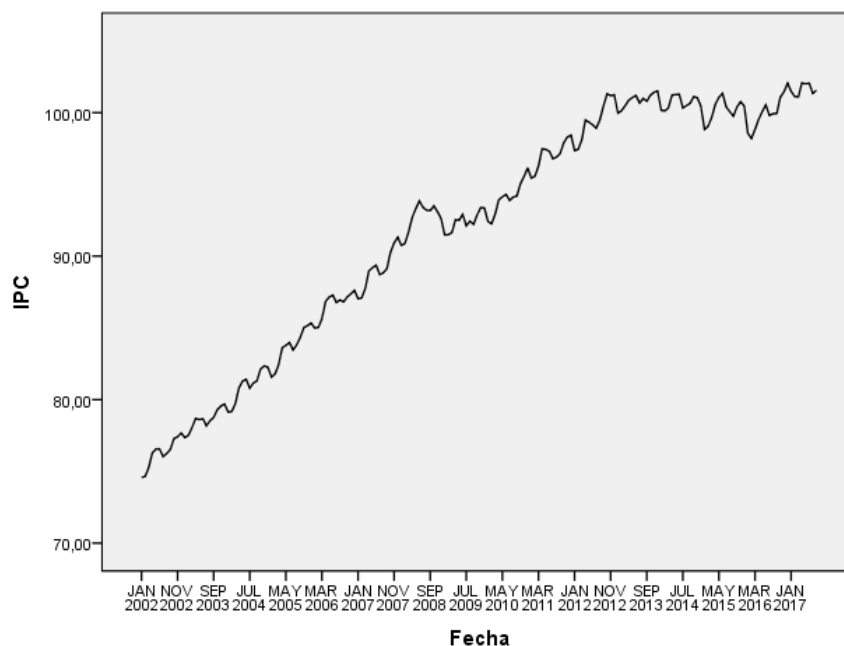


Figura 4.70. Gráfico de secuencia para la serie temporal del IPC nacional.

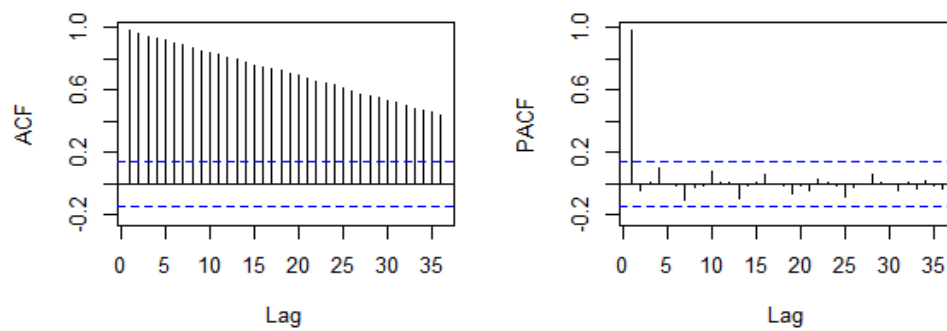


Figura 4.71. Gráfico de secuencia de la serie del IPC, junto a su FAS y FAP.

Se decide tomar una diferencia regular y otra estacional para transformar la serie en estacionaria. El resultado de la transformación sobre la serie se muestra en la imagen inferior. El test de Dickey-Fuller, con un P-valor menor que 0.01, confirma que la serie transformada es estacionaria.

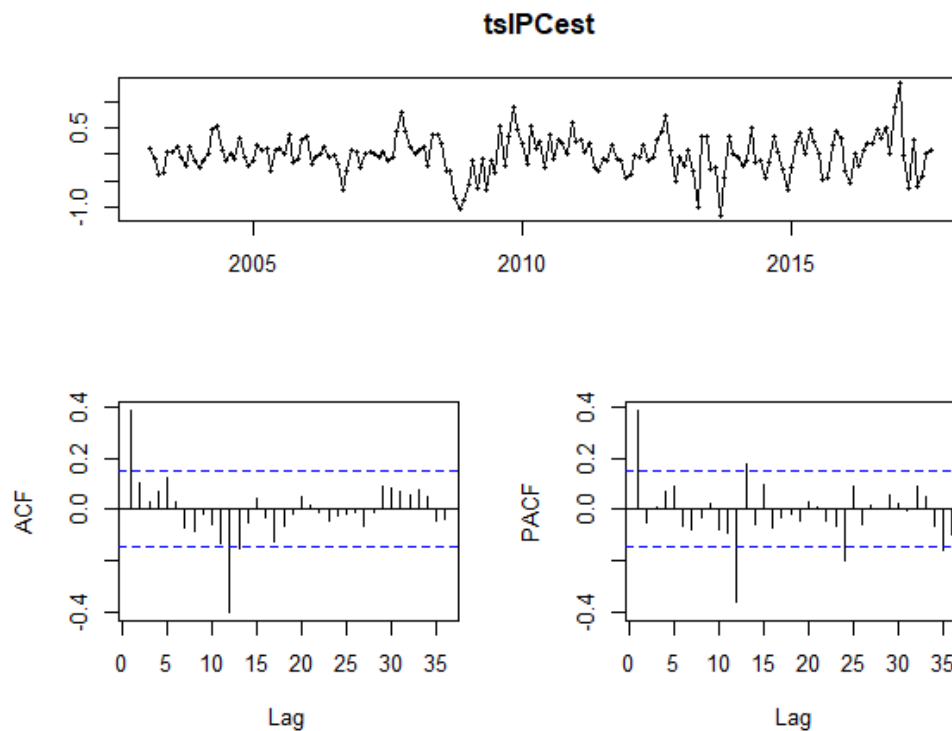


Figura 4.72. Serie del IPC transformada en estacionaria.

Se procede a ajustar un modelo ARIMA (0,1,1) x (0,1,1). Tras realizar el ajuste, la ecuación que explica el comportamiento de la serie es la siguiente:

$$\nabla \nabla_{12} z_t = (1 + 0.9076B^{12})(1 - 0.4513B)a_t \quad (4.13)$$

(-9.69) (6.87)

Una vez ajustado, se comprueba la bondad del ajuste estudiando los residuos generados. La imagen de la página siguiente muestra que la FAS y la FAP de los residuos no presentan palos que sobrepasen las bandas de confianza. Los test de Ljung-Box y Box-Pierce son superados con P-valores de 0.8357 y 0.837. Teniendo en cuenta que todos los parámetros del modelo resultan significativos, el ajuste se considera bueno.

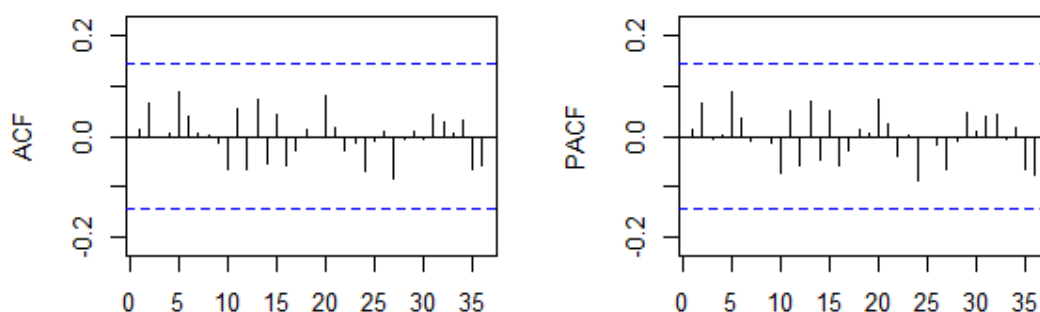


Figura 4.73. Residuos generados en el ajuste para la serie del IPC.

A continuación, se presenta una tabla recogiendo los atípicos detectados para esta serie:

Número	Tipo	Posición	Fecha	Magnitud	t
1	SLS	108	12/2010	0.2785	4.005

Tabla 4.15. Atípicos detectados para la serie del IPC.

El siguiente gráfico muestra el efecto del atípico sobre la serie del IPC:

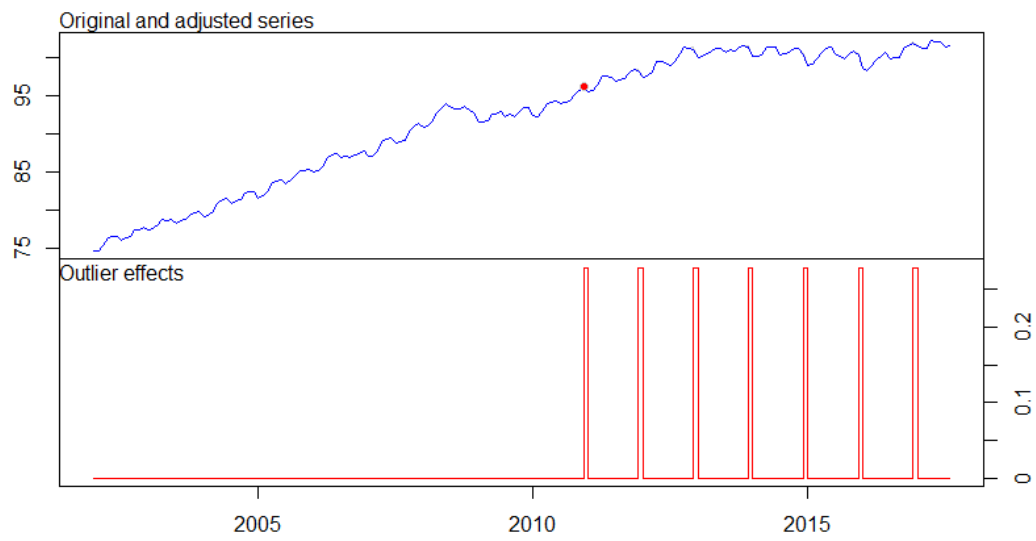


Figura 4.74. Efecto de los atípicos sobre la serie del IPC.

5 ANÁLISIS DE RESULTADOS.

En este último capítulo del trabajo, se expondrán las conclusiones obtenidas a raíz de analizar todas las series seleccionadas. El capítulo está estructurado en dos secciones: por un lado, se expondrán las conclusiones obtenidas sobre los efectos de la crisis en los diferentes sectores estudiados, así como la relación entre estos. Por otra parte, se demostrará la validez y utilidad de los modelos ARIMA a la hora de modelar el comportamiento de las series temporales y hacer predicciones futuras.

5.1 Impacto de la crisis en el sector de la construcción, el sector inmobiliario y la industria metalúrgica.

5.1.1 Tabla de outliers detectados en todas las series.

Se comienza esta sección mostrando una tabla a modo de recopilatorio de todos los atípicos detectados en los análisis sobre el total de las series temporales. Recordamos que los atípicos son valores de la serie que parecen ser generados de forma diferente al resto, sin seguir el modelo general de la serie. Establecer una recopilación de todos los atípicos detectados en un conjunto de series temporales, que se supone guardan algún tipo de relación entre sí, es una poderosa herramienta de análisis. Conocer la coincidencia de atípicos en un rango de tiempo, el desfase temporal entre ellos, el tipo y la magnitud, puede aportar información sobre como el comportamiento de unas series temporales puede afectar a otras. La tabla que se muestra a continuación recopila todos los atípicos detectados en el capítulo de análisis de series.

Año	Número	Atípicos
1997	2	S3-Dic-AO(+), S1-Dic-SAO(-),
1998	0	
1999	2	S1-Abr-LS(+), S11-Dic-AO(+)
2000	2	S3-Abr-AO(+), S11-Mar-AO(+)
2001	2	S1-Dic-SAO(-), S8-Dic-SLS(-)
2002	4	S3-Dic-AO(+), S1-Mar-AO(-), S1-Dic-AO(-), S12-feb-AO(+)
2003	4	S3-Dic-AO(+), S1-Dic-AO(-), S9-Jul-TC(-), S11-May-TC(+)
2004	1	S11-Nov-TC(+)

2005	5	S3-Mar-AO(+), S3-May-LS(+), S1-Abr-LT(+), S7-Sep-AO(+), S11-Dic-LS(+)
2006	2	S7-Mar-CT(+), S7-Abr-AO(-)
2007	3	S3-Dic-LT(-), S1-Mar-AO(+), S1-Jun-AO(+)
2008	16	S3-Jul-IO(-), S3-Ag-SAO(+), S1-Feb-LT(-), S1-Ag-SAO(+), S1-Dic-AO(-), S7-Mar-LS(-), S7-Abr-AO(+), S7-Dic-SLS(+), S6-Abr-AO(+), S5-Abr-AO(+), S8-Abr-TC(+), S8-jul-SLS(+), S8-Nov-LS(-), S9-Oct-LS(-), S11-Nov-LS(-), S12-Mar-TC(-)
2009	5	S3-En-TC(+), S3-May-LT(-), S1-May-LT(+), S9-Sep-LS(+), S10-Ene-TC(-)
2010	5	S1-Abr-LS(+), S1-Dic-SAO(+), S7-Ag-SLS(+), S6-Sep-LS(-), S13-Dic-SLS(+)
2011	2	S5-Ene-TC(+), S10-Mar-AO(+)
2012	1	S2-Sep-LS(+)
2013	1	S1-Mar-AO(-)

Tabla 5.1. Outliers detectados en el conjunto total de series analizadas.

En total se han detectado 57 outliers. A continuación, se muestra la evolución anual de los atípicos detectados entre 1997 y 2013:

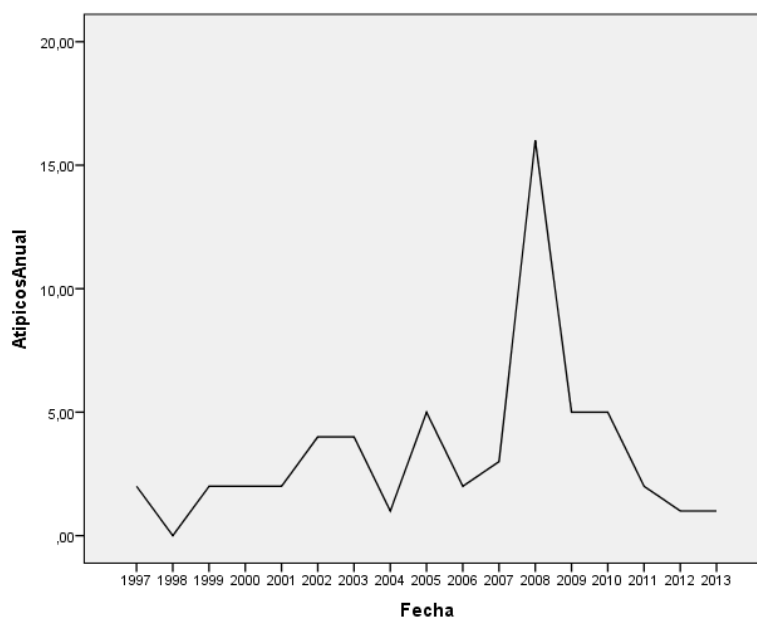


Figura 5.1. Evolución anual del número total de atípicos detectados.

Observando los resultados obtenidos, se confirma que el año 2008, el cual supuso el auge de la crisis económica en nuestro país, es el que acumula mayor cantidad de valores atípicos:

- El 28.07% de outliers se encuentra concentrado en el año 2008.
- Los años 2005, 2009 y 2010 presentan, cada uno, un 8.77% de concentración de atípicos.
- Los años 2003 y 2004 presentan, cada uno, un 7.01% del total de atípicos.

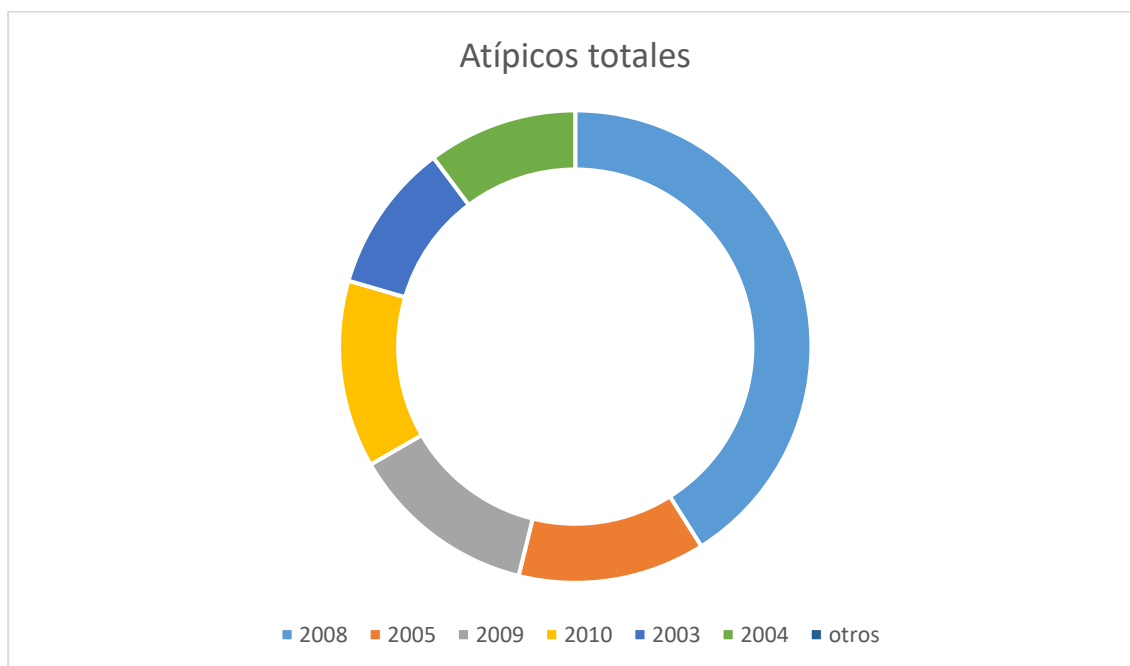


Figura 5.2. Atípicos totales detectados.

Si consideramos los años críticos de la crisis en nuestro país como 2008, 2009 y 2010, vemos que la acumulación total de atípicos en esos años alcanza el 45.61%

Una vez estudiada la distribución temporal de los atípicos, se procede a estudiar su origen. Las series analizadas pertenecen a cuatro grandes grupos: sector de la construcción, sector inmobiliario, industria metalúrgica e indicadores macroeconómicos generales. Respecto al año 2008, se pueden extraer las siguientes conclusiones:

- De los 16 atípicos detectados en el año 2008: el 31.25% pertenecen al sector de la construcción, otro 31.25% al sector inmobiliario, otro 31.25% a la industria metalúrgica, y el resto a los indicadores macroeconómicos. Se puede concluir que

la crisis del 2008 afectó de forma clara y generalizada a la actividad del sector de la construcción, el sector inmobiliario y la industria metalúrgica de igual forma.

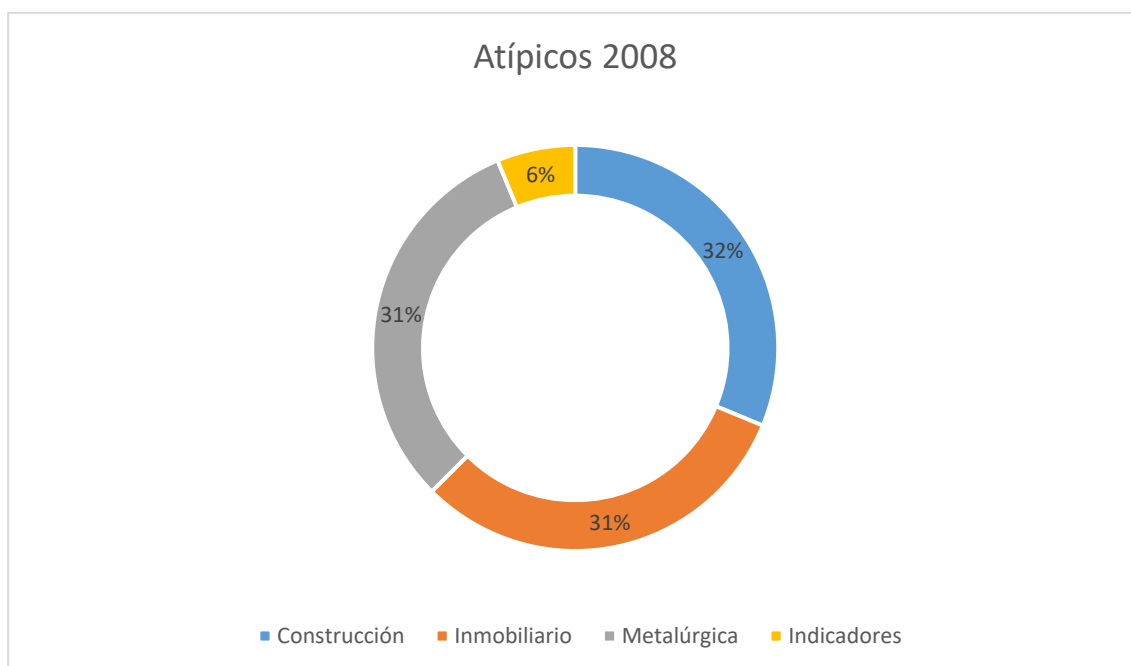


Figura 5.3. Distribución de los atípicos detectados en el año 2008.

- De todos los atípicos detectados para el año 2008: el 25% son atípicos aditivos (AO), otro 25% son cambios de nivel (LS), un 18.75% atípicos estacionales, un 12.5% son cambios transitorios (TC) y el resto se dividen entre otros tipos en porcentajes minoritarios. Esto quiere decir que, por lo general, la crisis actuó sobre estos sectores como fuertes impactos puntuales, o como cambios permanentes en el comportamiento de la serie.

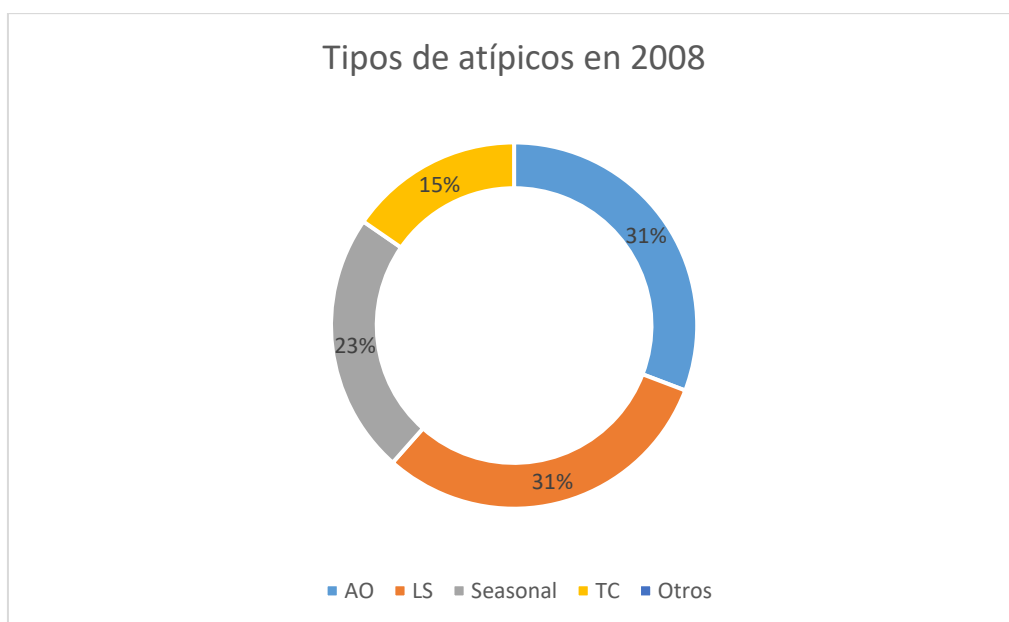


Figura 5.4. Tipos de atípicos detectados en 2008.

5.1.2 Impacto y efectos de la crisis sobre los sectores estudiados.

En este apartado, se analizará cronológicamente la aparición de los atípicos en torno a los años de la crisis. Mediante este análisis, se establecerá una posible cadena de causalidad, que relacionará la caída en los diferentes sectores estudiados. A continuación, se muestra una tabla que contiene los atípicos detectados en el año 2008 por cada mes del año.

Mes (2008)	Atípicos
Enero	-
Febrero	S1
Marzo	S7, S12
Abril	S7, S6, S5, S8
Mayo	
Junio	S8
Julio	S3
Agosto	S1, S3
Septiembre	
Octubre	S9
Noviembre	S11, S8
Diciembre	S1, S7

Tabla 5.2. Atípicos detectados en los meses del año 2008 (Se muestra solo el código de la serie).

Además de los hechos ocurridos en 2008, es importante incluir las observaciones realizadas durante el análisis de las series del sector inmobiliario. Recordando la serie del total de hipotecas constituidas en España, se observaba que ya a finales del 2005 la serie comenzaba a desacelerar su ascenso, y presentaba una meseta previa al colapso en 2008. Si observamos los atípicos detectados para esa misma serie, vemos que, en abril de 2006, justo dos años antes del periodo de mayor convulsión, la serie muestra un atípico aditivo negativo. Teniendo este atípico en cuenta, junto a la desaceleración temprana en la actividad del sector, unido al hecho de que el sector inmobiliario pertenece al sector servicios, siendo la “cúspide de la pirámide” desde la manipulación de la materia prima, pasando por la construcción como tal de viviendas, la siguiente hipótesis puede ser planteada:

En 2006, el sector inmobiliario comienza a desacelerarse y se estanca, fruto del inicio del colapso de la burbuja inmobiliaria. Entre este acontecimiento y 2008, se gesta la crisis económica en España. A comienzos de 2008, el sector de la construcción colapsa. Para este momento, el sector inmobiliario ya estaba sumido en la crisis. Algo más tarde, a finales de año, la industria metalúrgica sufre un descenso importante de actividad, sin llegar a un nivel de colapso tan gran como los otros dos sectores. Se puede ver como la crisis comienza en el sector servicios (sector inmobiliario), para poco a poco ir filtrándose al sector secundario (sector de la construcción), y finalmente afectar a la parte del sector secundario estrechamente relacionado con la manipulación y transformación de las materias primas (industria metalúrgica – acero).

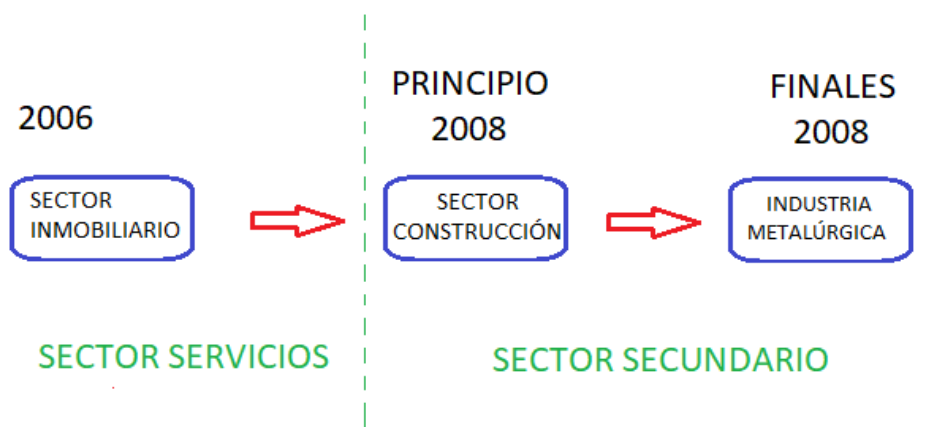


Figura 5.5. Filtración de la crisis a través de los sectores estudiados, reflejado en los atípicos detectados.

5.1.3 Posibles explicaciones y causas para los atípicos detectados.

En este apartado, se va a tratar de buscar acontecimientos concretos del panorama social y económico que puedan considerarse como causa directa de algunos de los atípicos que se han detectado, y reflejan el fenómeno de la crisis a través de los tres sectores estudiados.

Este proceso es especialmente importante, ya que de alguna forma cierra el círculo que representa el proceso seguido a lo largo de este trabajo, y en el análisis de series temporales: empezamos el estudio basándonos en datos extraídos de la realidad (las series temporales). A partir de esta medición de la realidad, se propusieron modelos matemáticos que la explicaran (ajustes de modelos ARIMA), y nos movimos en la dimensión matemática de estos modelos para establecer relaciones y causalidad entre las series estudiadas. En este último apartado, volvemos a la dimensión real con el fin de encontrar hechos reales detrás de las relaciones y posibles explicaciones que hemos encontrado en el dominio matemático.

A continuación, se presentan algunos acontecimientos que coinciden con las fechas de aparición de atípicos en las series analizadas, y que pueden ser causas explicativas de los mismos;

- En **abril de 2006** se percibió un atípico aditivo negativo en el número total de hipotecas registradas. Durante esos meses, el sector inmobiliario vivía los últimos momentos de su época dorada en nuestro país, mostrando un pico de actividad. Es precisamente en abril de 2006 cuando Renta Corporación efectúa su salida a bolsa. Aunque su estreno en bolsa supuso un éxito, con un aumento del 5.22%, no lo fue tanto si se tiene en cuenta que se esperaba que este aumento fuese mayor del 10%. Este aumento menor de lo esperado se debió al rápido desinflamiento de las acciones después de la subida inicial. Este importante desvío negativo en las predicciones (cercano al 50%) no puede ser pasado por alto, ¿Cómo pudo alejarse tanto la realidad de las predicciones, siendo la empresa una importante entidad de un sector en alza, que producía una enorme rentabilidad económica? Durante meses, el sector, los inversores y las autoridades bursátiles, ignoraron hechos como este, confiando en la robustez del sector y su aparentemente infinita proyección económica. No obstante, como quedó demostrado pocos años después, esta falta de previsión permitió a la burbuja estallar con más fuerza, lo que desembocó en una crisis económica en

España notablemente más recrudecida que en el resto de los países occidentales desarrollados.

- En **marzo de 2008** el Euribor cierra al 4.59%, situándose 0.241 puntos por encima del mes anterior, y 0.484 que en marzo de 2007. Esta subida explica el atípico negativo detectado en la serie del número de hipotecas totales constituidas en España. Además, el vicepresidente económico Pedro Solbes anuncio una de las primeras reducciones en la previsión del crecimiento en PIB, alegando la crisis del sector inmobiliario como factor determinante. Esta reducción instaura un sentimiento negativo en el panorama económico español. Una reacción típica en estas situaciones por parte de las empresas consiste en acercarse a posturas conservadores, apostando por minimizar perdidas. Por lo tanto, este evento puede explicar el atípico negativo detectado en la serie del parto total registrado en España.
- En **abril de 2008** ocurren un hecho que dio un leve respiro a la economía española. El 18 de abril, el gobierno aprueba el un plan de 10.000 millones de euros para reactivar la economía española. Este hecho explica por qué se registran atípicos aditivos positivos en las series de las hipotecas totales registradas, la compraventa de viviendas tanto nuevas como de segunda mano, y el IPI de la industria metalúrgica.
- En **julio de 2008**, Martinsa Fadesa (sector inmobiliario) presenta el mayor concurso de acreedores jamás registrado en nuestro país, por una deuda superior a 8000 millones de euros (proceso que aun a día de hoy no ha concluido). Además, el presidente del gobierno, José Luis Rodríguez Zapatero, habla de “crisis” económica por primera vez.
- En **agosto de 2008**, el gobierno aprueba un plan de choque con 24 medidas para detener la crisis. Este hecho coincide con un atípico positivo en el número de asociados a la seguridad social en el sector de la construcción.
- En **octubre de 2008**, la bolsa pierde un 8.16%, siendo la segunda mayor caída histórica registrada. En este punto la crisis se ha desarrollado en todos los ámbitos de la economía española, que empieza a colapsar. Se empiezan a notar los efectos en la industria, como la metalúrgica, y se detecta un cambio de nivel negativo en la producción interna de acero. Un mes más tarde, se detecta el mismo fenómeno en la serie de las importaciones de acero.

- En **diciembre de 2008**, España entra oficialmente en recesión. La serie del número de asociados a la seguridad social en el sector de la construcción presenta un nuevo atípico aditivo negativo.

A raíz de los hechos que se acaban de exponer, se demuestra que la hipótesis realizada sobre el efecto de la crisis y su avance en base a los atípicos detectados, es respaldada en varios puntos del tiempo por acontecimientos reales dentro del proceso de la crisis que explican estos atípicos.

En este apartado, ha quedado confirmado la alta capacidad de los modelos ARIMA como herramienta de estudio y diagnóstico. A través de análisis de series temporales y la detección de outliers, se ha podido comprender la propagación temporal de la crisis en los sectores estudiados. En el próximo apartado, se complementará este al demostrad la potencia de los modelos ARIMA como herramienta de predicción.

5.2 Predicciones con modelos ARIMA.

En este apartado, se va a tratar de demostrar la efectividad de los modelos ARIMA a la hora de realizar predicciones a futuro fiables. Para ello, se va a tomar una de las series estudiadas, y se va a interrumpir a lo largo de varios puntos que supongan un cambio drástico en el comportamiento de la serie. Para cada uno de esos cortes, se van a realizar predicciones en base a un modelo ARIMA, que después se van a comparar con los datos reales. Por último, se utilizará el modelo que se ajustó para serie y se harán predicciones con él. Estas predicciones se compararán con los datos más recientes publicados que no estaban disponibles a la hora de realizar el trabajo.

La serie temporal elegida para este estudio es la de trabajadores asociados a la seguridad social en el sector de la construcción. Las predicciones se realizarán sobre los siguientes cuatro puntos en el tiempo:

- Diciembre de 2002
- Marzo de 2007
- Abril de 2010
- A partir del final de la serie en septiembre de 2017

Se han elegido estos puntos porque, salvo el punto final, en todas estas fechas se detectaron valores atípicos. La siguiente figura muestra la serie temporal de los asociados a la seguridad social en el sector de la construcción, con los cortes indicados:

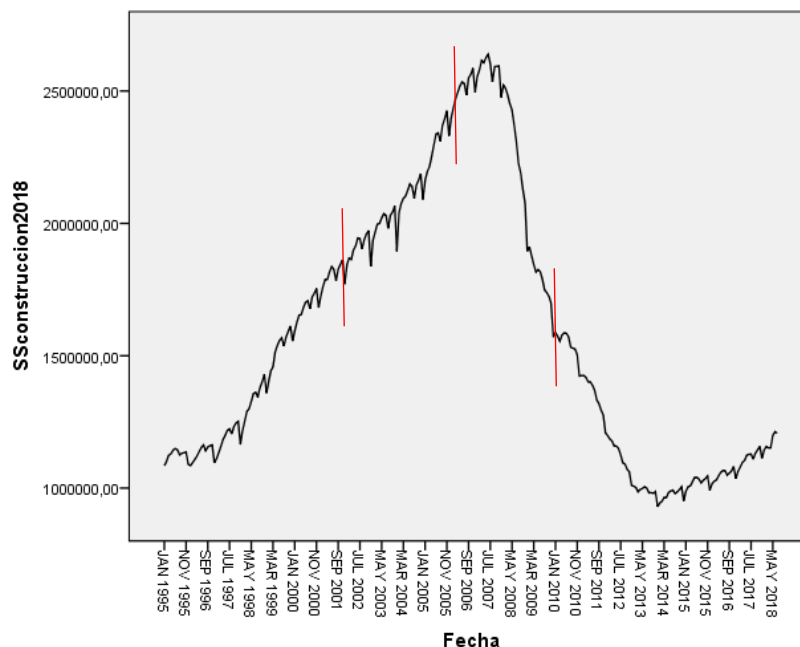


Figura 5.6. Serie de los asociados a la seguridad social en el sector de la construcción con los 3 cortes indicados.

5.2.1. Diciembre de 2002.

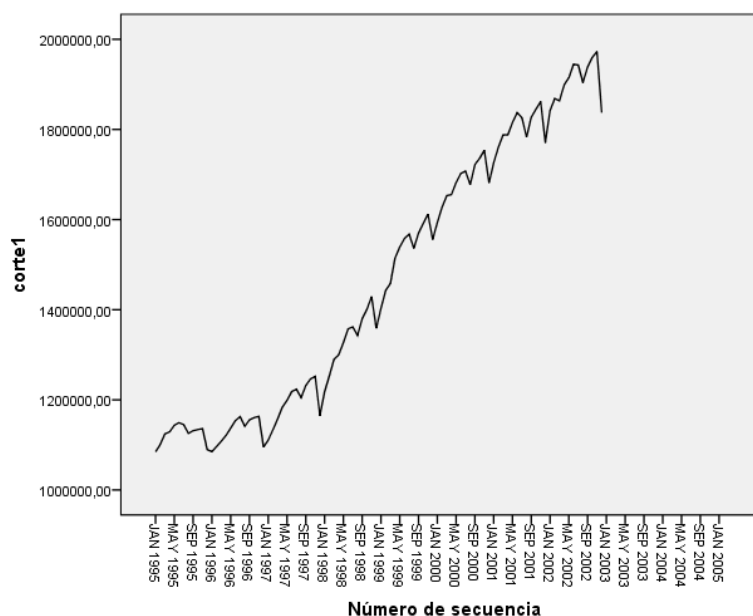


Figura 5.7. Serie de los asociados a la seguridad social en el sector de la construcción desde enero de 1996 hasta diciembre de 2002.

Para realizar la primera predicción, se usan los datos de la serie desde enero de 1995 hasta diciembre de 2002. Como el conjunto de datos a cambiado, será necesario ajustar un modelo ARIMA distinto al de la serie total. Este ajuste se realiza con la misma metodología vista en el apartado de análisis de las series estudiadas. Debido al reducido tamaño de la muestra, el modelo ajustado es extremadamente simple, siendo un ARIMA (0,1,0) x (0,1,1):

$$\nabla \nabla_{12} z_t = (1 - 0.407B^{12})a_t \quad (5.1)$$

(3.405)

Estadísticos del modelo

Modelo	Número de predictores	Estadísticos de ajuste del modelo	Ljung-Box Q(18)			Número de valores atípicos
		R cuadrado estacionaria	Estadísticos	GL	Sig.	
corte1-Modelo_1	0	,699	7,716	17	,972	6

Parámetros del modelo ARIMA

				Estimación	SE	t	Sig.
corte1-Modelo_1	corte1	Sin transformación	Diferencia	1			
			Diferencia estacional	1			
			MA, estacional Retardo 1	,407	,120	3,405	,001

Valores atípicos

			Estimación	SE	t	Sig.
corte1-Modelo_1	Ene 1997	Tendencia local	10416,750	2324,787	4,481	,000
	Dic 1997	Aditivo estacional	-32924,492	5760,631	-5,715	,000
	Abr 1999	Cambio de nivel	47818,668	7197,840	6,643	,000
	Dic 2001	Aditivo estacional	-25229,157	6054,820	-4,167	,000
	Mar 2002	Aditivo	-31946,524	6118,303	-5,221	,000
	Dic 2002	Aditivo	-42937,946	8912,086	-4,818	,000

Tabla 5.3. Resultados totales para el ajuste del modelo del primer corte.

A continuación, se muestra el gráfico de secuencia de la serie cortada junto a la predicción de los valores hasta abril de 2005. Además, se incluyen los límites de confianza superior e inferior.

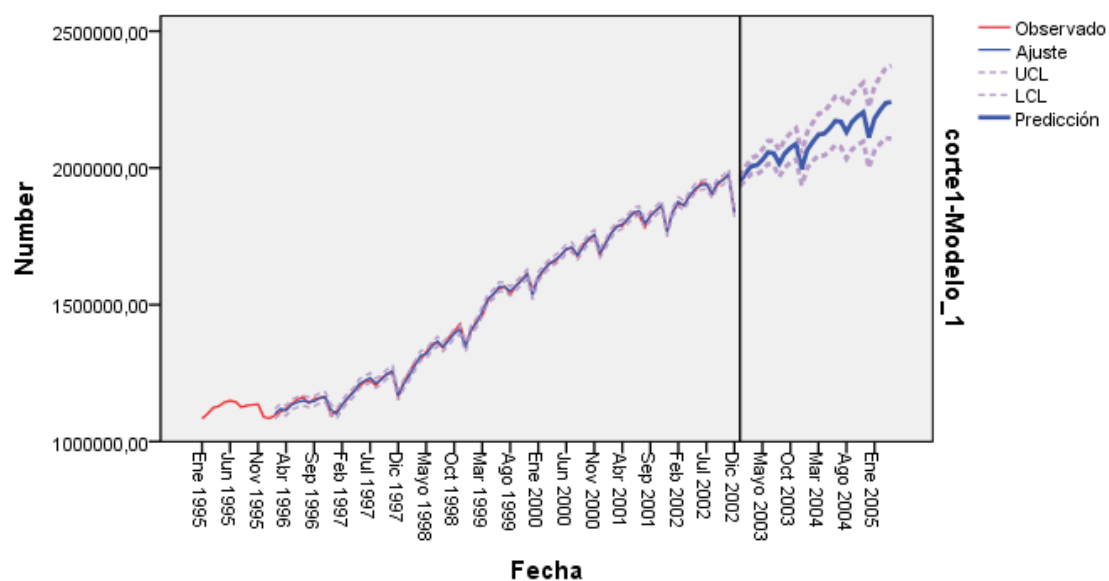


Figura 5.8. Gráfico de secuencia para las predicciones a partir de diciembre de 2002, junto a sus límites de confianza superior e inferior.

A continuación, se muestra un gráfico comparativo entre la serie original superpuesta con la serie completada con las predicciones:

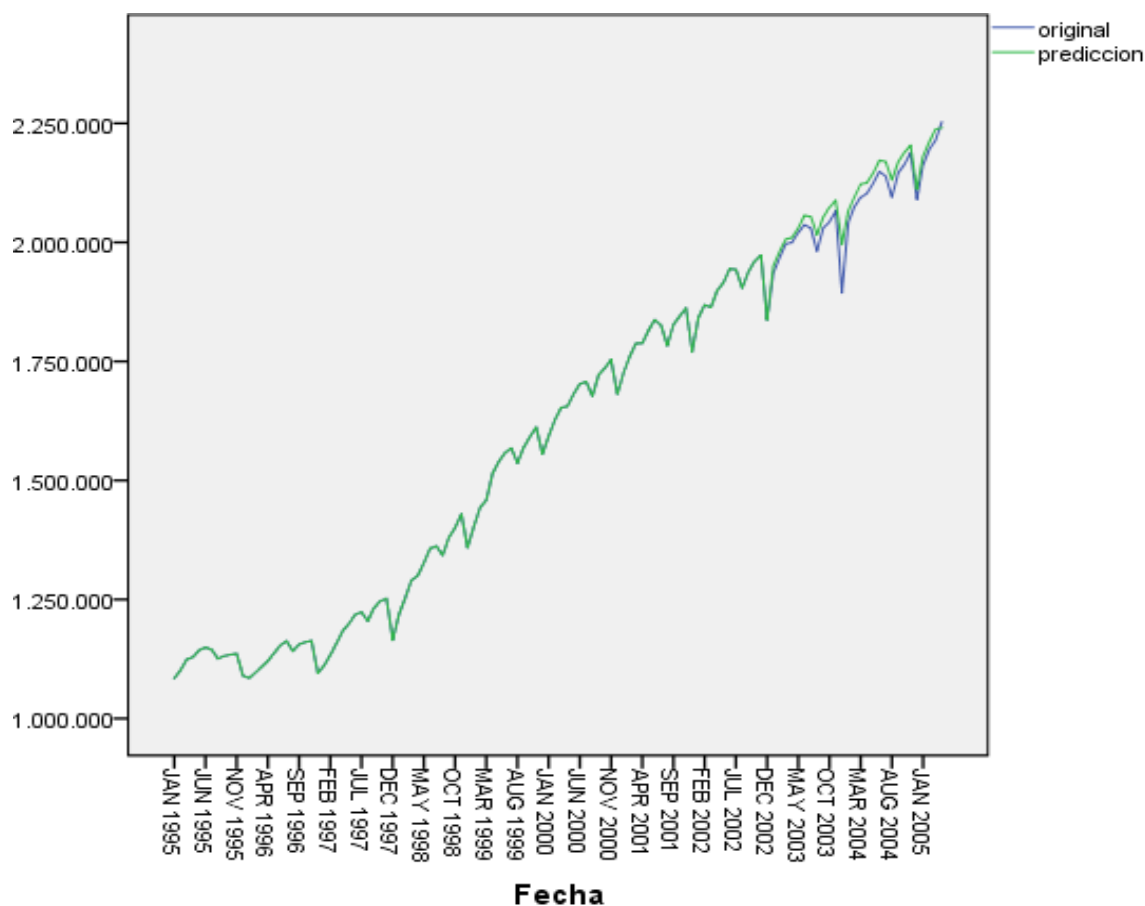


Figura 5.9. Gráfico de secuencia para las predicciones a partir de diciembre de 2002 en contraposición a los datos reales.

A simple vista, parece que las predicciones presentan un alto grado de coincidencia con la realidad. Este resultado es esperable dado el intervalo temporal estudiado, caracterizado por un comportamiento y tendencia estables, así como un ciclo estacional bien definido y repetitivo. No obstante, a continuación, se presenta una tabla con los límites de confianza superiores e inferiores obtenidos en las predicciones (enero 2003-abril 2005), acompañados por los valores reales. De esta forma, se puede estudiar el nivel de exactitud de las predicciones con respecto a los datos reales. La primera fila de la tabla corresponde a la predicción de enero de 2003 y la última a abril de 2005. Por motivos de espacio, las fechas no pueden ser incluidas como una quinta columna.

Límite inferior predicción	Datos reales	Límite superior predicción	Predicciones
1933600,04	1936072,00	1967683,75	1950641,89
1956373,35	1967322,00	2004574,96	1980474,16
1977393,42	1997346,00	2036428,09	2006910,76
1975790,64	1999484,00	2043958,00	2009874,32
1992098,11	2020684,00	2068311,53	2030204,82
2014968,49	2036719,00	2098456,11	2056712,30
2009014,21	2030019,00	2099191,14	2054102,67
1967610,95	1981238,00	2064014,14	2015812,54
2002392,53	2030445,00	2104643,55	2053518,04
2019539,99	2042780,00	2127322,03	2073431,01
2031798,47	2065942,00	2144841,23	2088319,85
1936515,14	1894249,00	2054584,45	1995549,79
2001036,45	2040879,00	2130989,04	2066012,74
2025426,65	2075220,00	2166263,36	2095845,00
2046812,60	2094439,00	2197750,61	2122281,60
2045042,94	2102965,00	2205447,40	2125245,17
2060904,40	2124209,00	2230246,94	2145575,67
2083167,17	2148781,00	2260999,12	2172083,15
2076506,44	2138973,00	2262440,60	2169473,52
2034334,52	2094834,00	2228032,26	2131183,39

2068307,92	2145544,00	2269469,86	2168888,89
2084622,41	2163584,00	2292981,30	2188801,85
2096032,98	2187648,00	2311348,41	2203690,70
1999893,57	2090195,00	2221947,71	2110920,64
2064275,22	2162575,00	2298491,96	2181383,59
2088326,76	2195708,00	2334104,93	2211215,85
2109242,62	2214051,00	2366062,28	2237652,45
2106913,19	2252430,00	2374318,83	2240616,01

Tabla 5.4. Predicciones sobre diciembre de 2002 comparadas con los límites superior e inferior de confianza.

A raíz de los resultados obtenidos, vemos que comparando los datos reales con los límites superior e inferior de confianza, solo uno de los datos reales queda fuera de dichos niveles. Esto se traduce en que para los niveles de confianza establecidos, ha habido un 96.4% de acierto. Este porcentaje entra dentro del nivel de confianza del 95% con el que se trabaja, por lo que las predicciones se pueden considerar en un éxito.

5.2.2. Marzo de 2007.

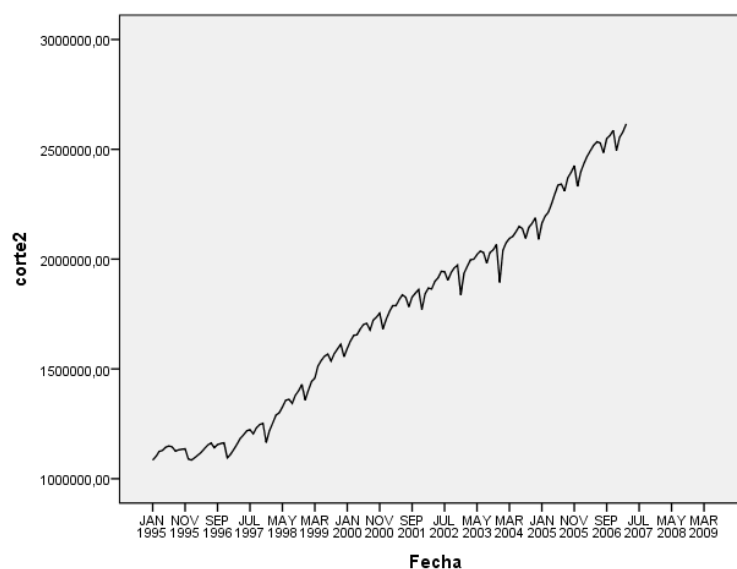


Figura 5.10. Serie de asociados a lo seguridad social en el sector de la construcción desde enero de 1996 hasta marzo de 2007.

La predicción sobre marzo de 2007 se realiza de la misma forma que la predicción anterior. En este caso se ajusta un modelo ARIMA (1,1,0) x (0,1,1):

$$(1 - 0,172B)\nabla\nabla_{12}z_t = (1 - 0,476B^{12})a_t \quad (5.2)$$

(2.04)

(5.661)

Estadísticos del modelo

Modelo	Número de predictores	Estadísticos de ajuste del modelo	Ljung-Box Q(18)			Número de valores atípicos
		R cuadrado estacionaria	Estadísticos	GL	Sig.	
corte2-Modelo_1	0	,781	13,293	16	,651	8

Parámetros del modelo ARIMA

					Estimación	SE	t	Sig.
corte2-Modelo_1	corte2	Sin transformación	AR	Retardo 1	,172	,086	2,014	,046
			Diferencia		1			
			Diferencia estacional		1			
			MA, estacional	Retardo 1	,476	,084	5,661	,000

Valores atípicos

			Estimación	SE	t	Sig.
corte2-Modelo_1	Dic 1996	Aditivo estacional	-28797,797	5509,754	-5,227	,000
	Ene 1998	Tendencia local	10513,130	2630,919	3,996	,000
	Abr 1999	Cambio de nivel	48694,048	7329,817	6,643	,000
	Dic 2001	Aditivo estacional	-26884,037	5099,732	-5,272	,000
	Mar 2002	Aditivo	-31418,786	4764,663	-6,594	,000
	Dic 2002	Aditivo	-34673,806	5112,389	-6,782	,000
	Dic 2003	Aditivo	-76950,982	4997,976	-15,396	,000
	Abr 2005	Innovador	33426,751	8566,559	3,902	,000

Tabla 5.5. Resultados globales para el ajuste del modelo del segundo corte.

A continuación, se muestra el gráfico de secuencia de las predicciones realizadas a partir de marzo de 2007, junto a sus intervalos de confianza:

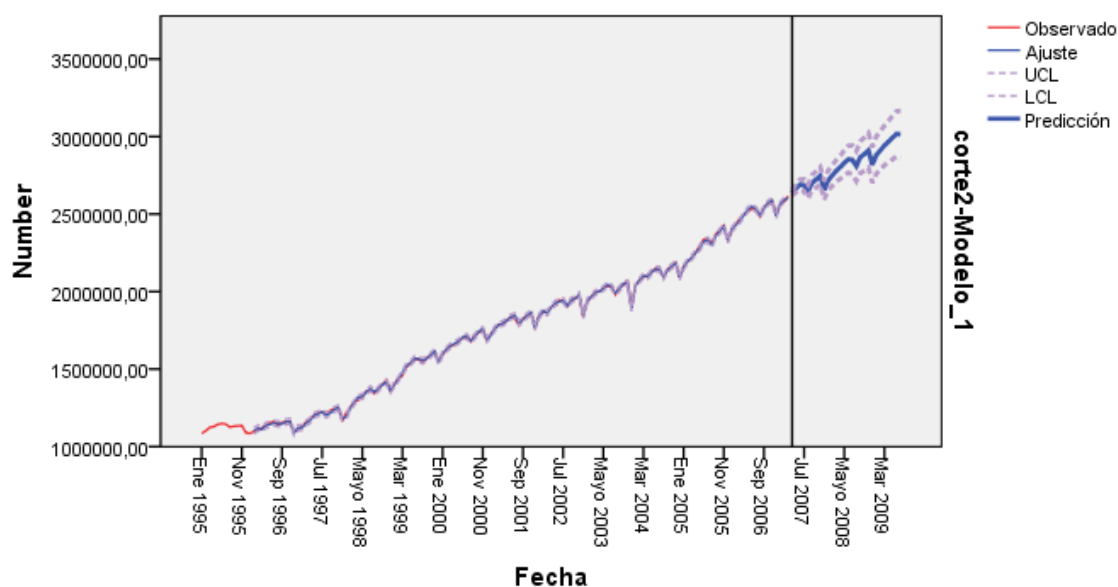


Figura 5.11. Gráfico de secuencia para las predicciones a partir de marzo de 2007, junto a sus límites de confianza superior e inferior.

En el siguiente gráfico podemos ver estas predicciones comparadas directamente con la serie original entre enero de 1995 y julio de 2009:

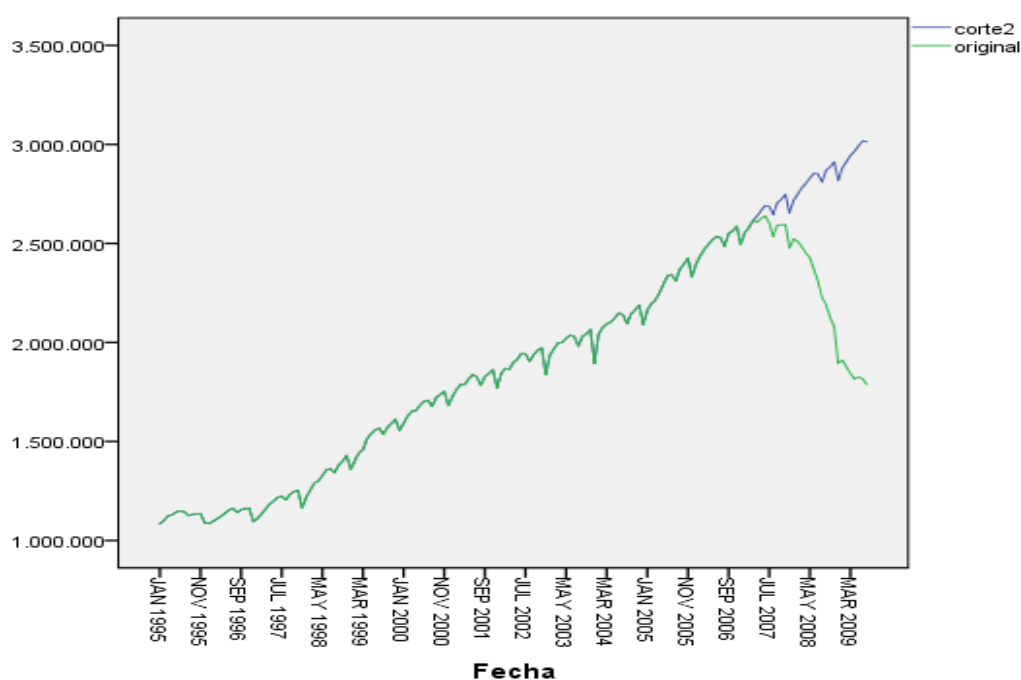


Figura 5.12. Gráfico de secuencia para las predicciones a partir de marzo de 2007 en contraposición a los datos reales.

En este caso las predicciones se alejan totalmente de la realidad, debido al impacto de la crisis sobre el que se habló en el apartado anterior. No tiene sentido comparar numéricamente

las predicciones con los datos, ya que a simple vista se ve que no habría ningún tipo de similitud. Aunque esta falta de coincidencia es debida en gran parte a que se ha realizado este segundo corte de tal forma que así ocurra, es un claro ejemplo de cómo un impacto de enorme poder como la crisis económica puede cambiarlo todo en cuestión de dos meses. Por lo tanto, aunque se disponga de un modelo explicativo y satisfactorio, es necesario actualizarlo periódicamente. Aun así, habrá casos extremos como el que está aquí representado en el que será necesario complementar los modelos ARIMA con un profundo estudio de la realidad para evitar consecuencias catastróficas derivadas de obtener predicciones completamente opuestas a la realidad, como ocurre en este caso.

5.2.3. Abril de 2010.

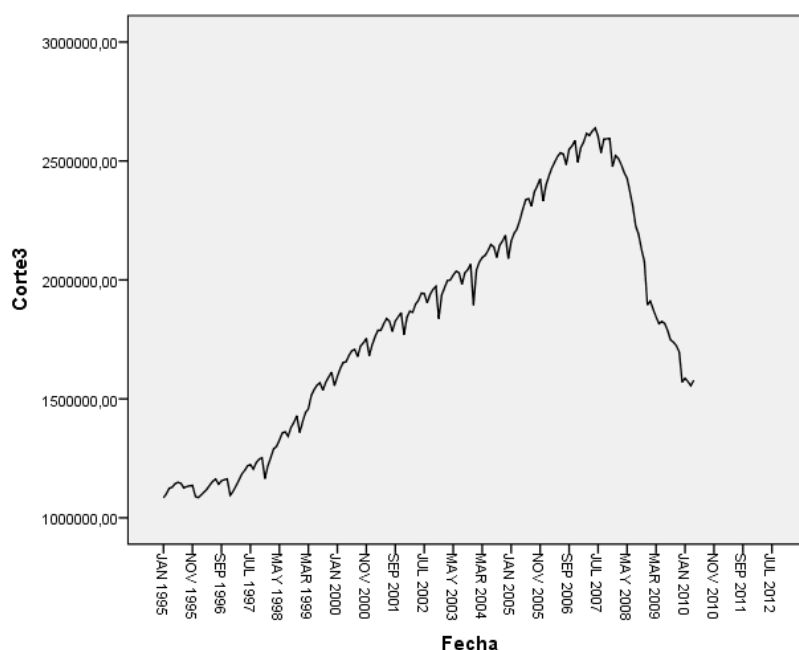


Figura 5.13. Serie del número de asociados a la seguridad social desde enero de 1996 hasta abril de 2010.

La predicción sobre abril de 2010 se realiza de la misma forma que las dos predicciones anteriores. En este caso se ajusta un modelo ARIMA (0,2,1) x (0,1,1):

$$\nabla^2 \nabla_{12} z_t = (1 - 0.354B^{12})(1 - 0.850B)a_t \quad (5.3)$$

(3.831) (18.652)

En la imagen inferior se pueden comprobar los resultados totales obtenidos para el ajuste de este modelo.

Estadísticos del modelo

Modelo	Número de predictores	Estadísticos de ajuste del modelo	Ljung-Box Q(18)			Número de valores atípicos
		R cuadrado estacionaria	Estadísticos	GL	Sig.	
Corte3-Modelo_1	0	,900	10,006	16	,866	13

Valores atípicos

			Estimación	SE	t	Sig.
Corte3-Modelo_1	Dic 1996	Aditivo estacional	-24063,902	6138,725	-3,920	,000
	Dic 1997	Aditivo	-20983,197	5096,583	-4,117	,000
	Abr 1999	Cambio de nivel	46799,034	7244,467	6,460	,000
	Dic 2001	Aditivo estacional	-26483,745	5869,205	-4,512	,000
	Mar 2002	Aditivo	-31339,402	4971,950	-6,303	,000
	Dic 2002	Aditivo	-34951,360	5392,998	-6,481	,000
	Dic 2003	Aditivo	-76462,987	5294,729	-14,441	,000
	Abr 2005	Tendencia local	18724,381	4318,977	4,335	,000
	Mar 2007	Aditivo	21074,928	4964,614	4,245	,000
	Feb 2008	Tendencia local	-36246,826	4619,525	-7,846	,000
	Ago 2008	Aditivo estacional	38227,719	5857,009	6,527	,000
	Dic 2008	Aditivo	-21544,528	5071,069	-4,249	,000
	Mayo 2009	Tendencia local	44576,483	5276,272	8,448	,000

Parámetros del modelo ARIMA

				Estimación	SE	t	Sig.
Corte3-Modelo_1	Corte3	Sin transformación	Diferencia	2			
		MA	Retardo 1	,850	,046	18,652	,000
		Diferencia estacional		1			
		MA, estacional	Retardo 1	,354	,092	3,831	,000

Tabla 5.6. Resultados totales para el tercer y último corte de la serie de asociados a la seguridad social en el sector de la construcción.

El siguiente gráfico muestra las predicciones realizadas en base al modelo ARIMA ajustado para el tramo de la serie bajo estudio:

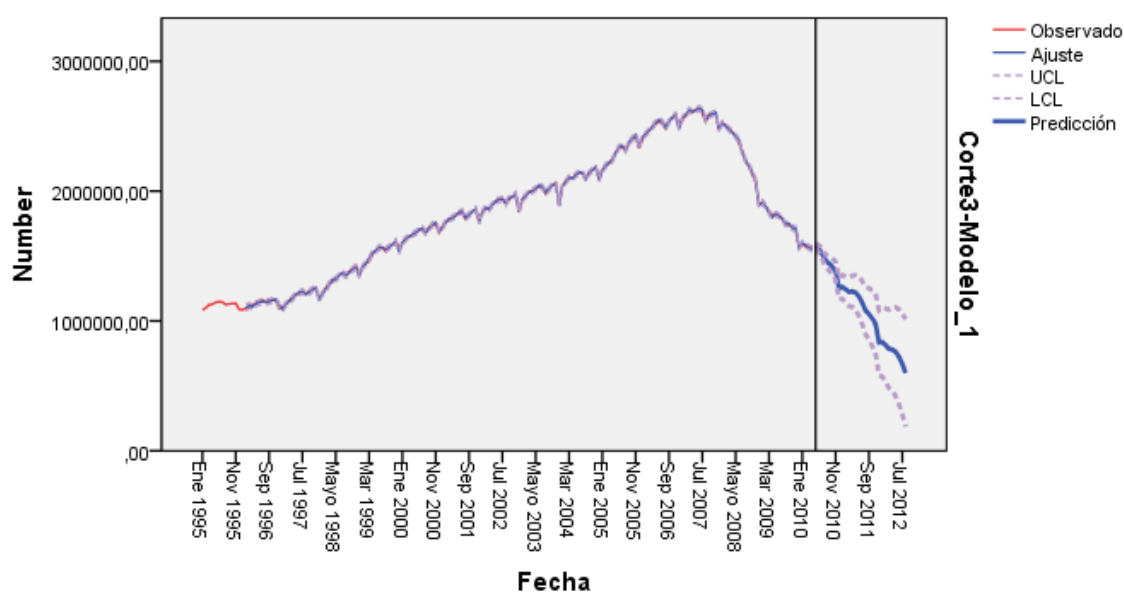


Figura 5.14. Gráfico de secuencia para las predicciones a partir de abril de 2010, junto a sus límites de confianza superior e inferior.

A continuación, se muestra un gráfico donde se comparan los datos reales con las predicciones realizadas a partir de abril de 2010. En este caso, se incluye también la representación gráfica de los límites inferior y superior de confianza, para facilitar el análisis visual.

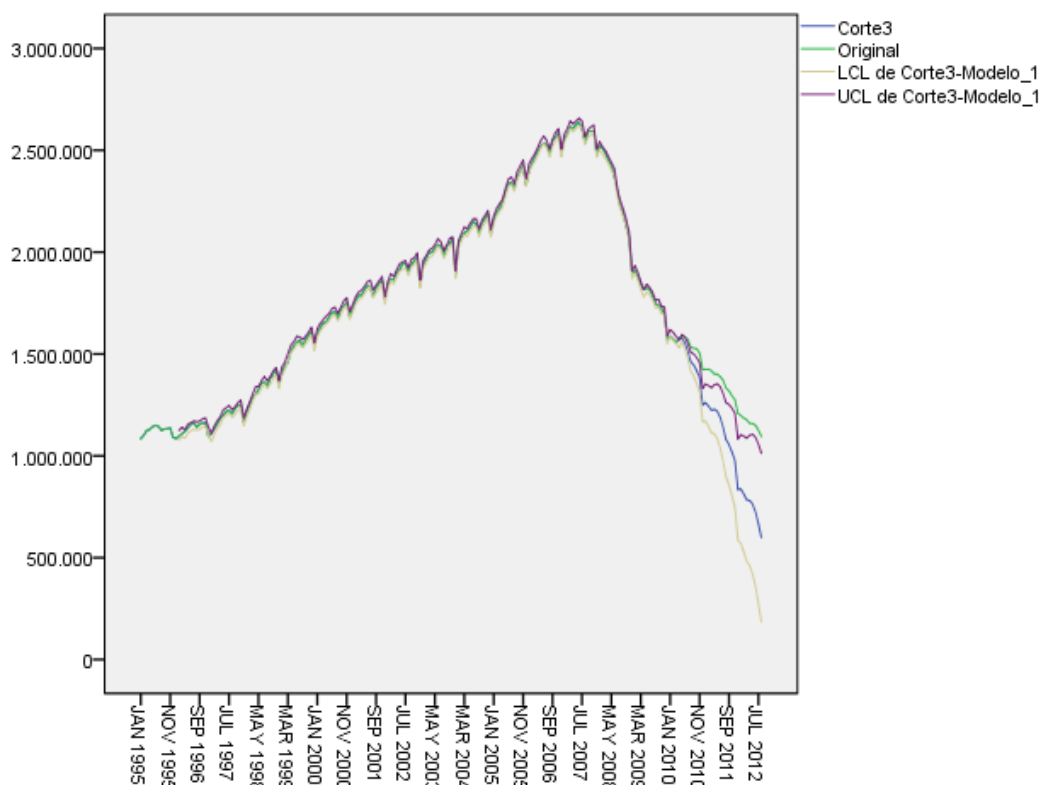


Figura 5.15. Gráfico de secuencia para las predicciones a partir de abril de 2010, junto a sus límites de confianza superior e inferior, y contrapuestos con los datos reales.

Para este punto, se pueden aplicar las mismas conclusiones que para el anterior. Las predicciones auguran un descenso en el número de asociados a la seguridad social en el sector de la construcción notablemente superior a lo que los datos reflejan. No obstante, la discrepancia entre ambos es menor que en el punto anterior, por lo que la inclusión de los límites de confianza es importante para determinar a simple vista que efectivamente los datos reales quedan completamente fuera del intervalo de confianza de las predicciones.

5.2.4. Septiembre de 2017

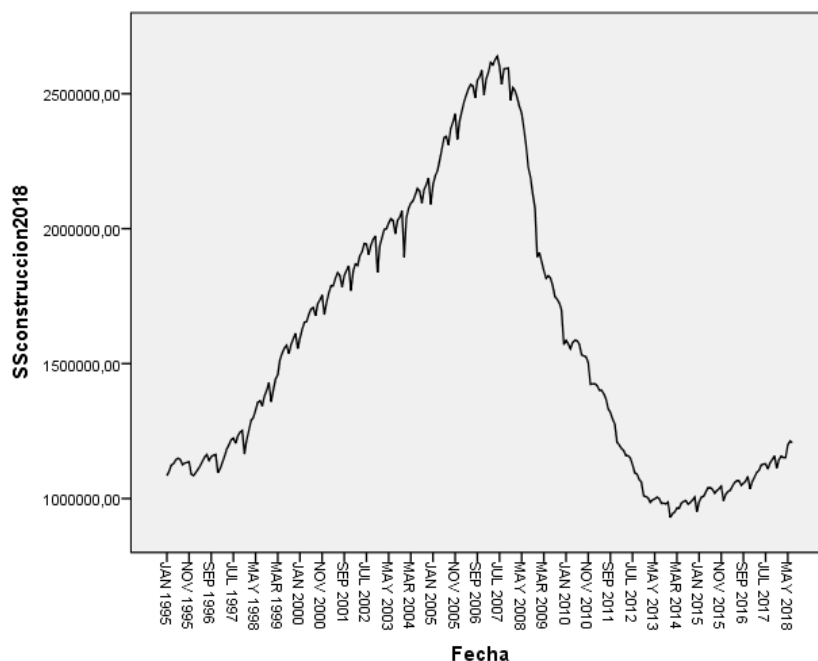


Figura 5.16. Serie de asociados a la seguridad social en el sector de la construcción.

En este último punto se utilizará el modelo obtenido en el apartado de análisis de las series para la serie estudiada. En base a este modelo ARIMA que se obtuvo previamente, se harán predicciones que se cotejarán con los datos reales disponibles en este momento, y que no lo estaban en el momento de comenzar a realizar este trabajo. De esta forma, se comprobará de forma efectiva la eficacia y utilidad de los modelos ARIMA obtenidos a lo largo de todo el trabajo a la hora de predecir el comportamiento de las series temporales. En el gráfico inferior se muestran las predicciones generadas por el modelo a partir de septiembre de 2017, llegando hasta julio de 2018.

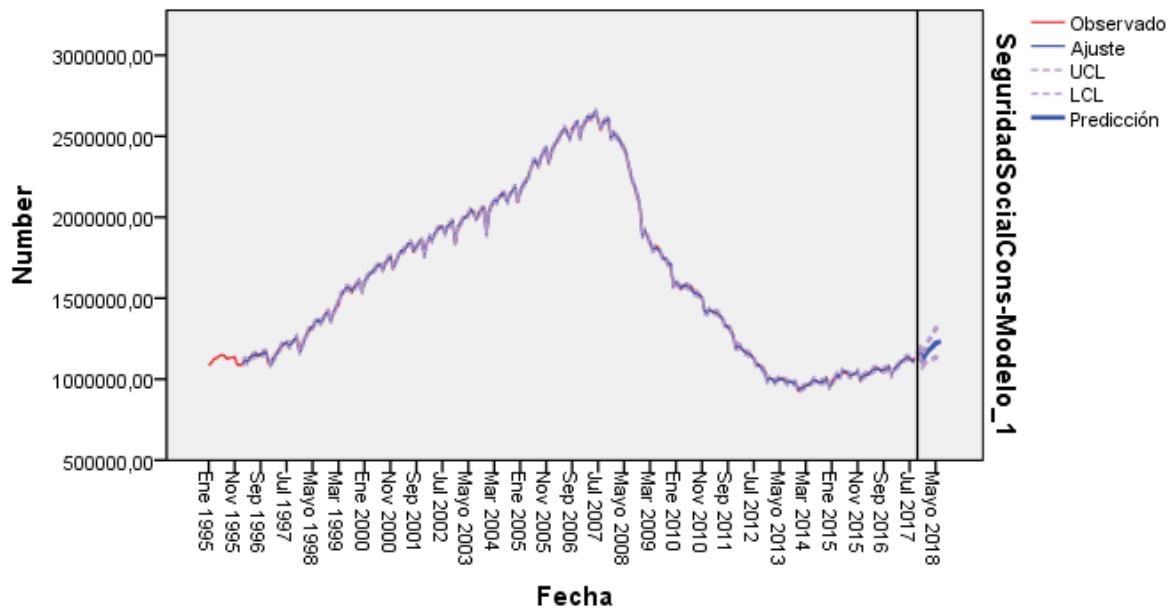


Figura 5.17. Predicciones para la serie de los asociados a la seguridad social.

A continuación, se facilita el gráfico que compara las predicciones con los nuevos datos obtenidos para el periodo actual.

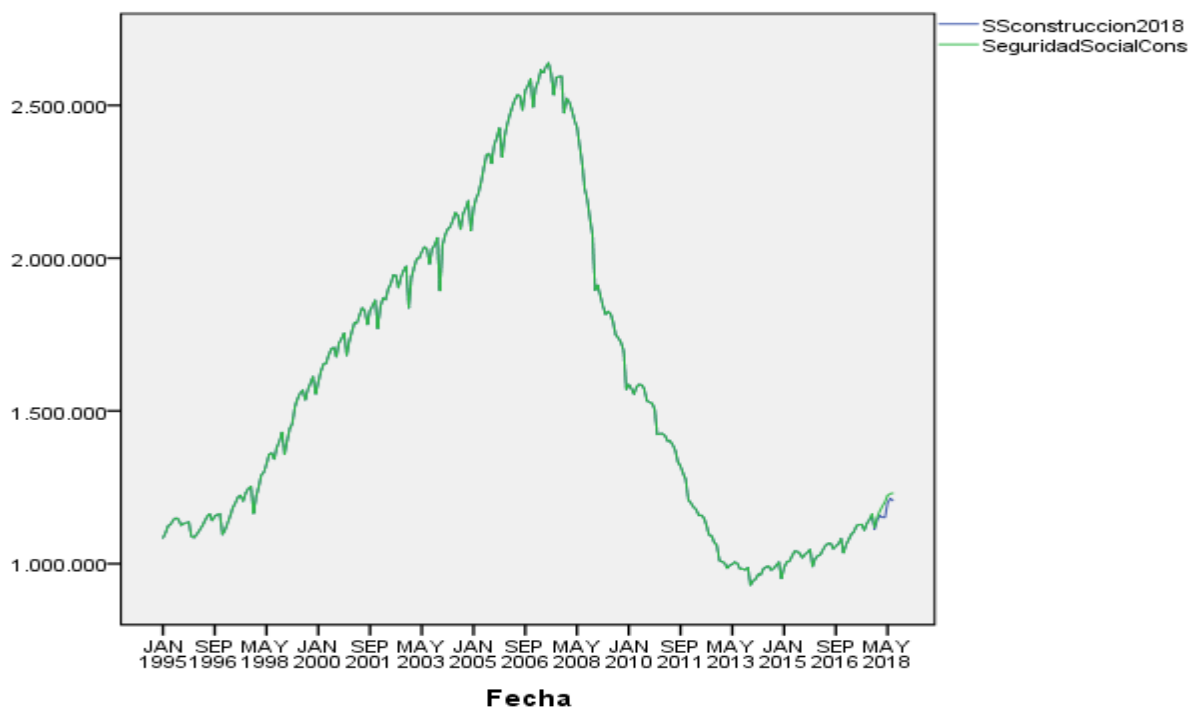


Figura 5.18. Predicciones para la serie de la seguridad social comparados con los nuevos datos publicados.

En este caso, al tratarse de pocos datos, es difícil determinar a simple vista la exactitud de las predicciones generadas por el modelo. Por ello, se elabora la siguiente tabla donde se recogen los datos reales, las predicciones, y los límites de confianza de estas últimas.

Límite inferior	Datos	Límite superior	Predicciones
1127020,12	1144608,00	1161305,21	1117061,73
1135744,78	1157481,00	1190195,82	1144162,67
1082169,65	1113240,00	1156124,69	1162970,30
1105530,37	1143680,00	1199221,91	1119147,17
1115722,82	1156840,00	1229607,91	1152376,14
1121394,52	1151650,00	1255980,05	1172665,36
1124142,47	1151650,00	1279925,16	1188687,29
1135072,54	1198748,00	1312516,26	1202033,82
1129415,27	1213337,00	1328943,01	1223794,40
1119378,34	1207035,00	1341370,30	1229179,14

Tabla 5.7. Predicciones sobre septiembre de 2017 y datos reales, comparadas con los límites superior e inferior de confianza.

En una primera inspección, vemos que todos los datos están incluidos dentro de los límites de confianza de las predicciones, por lo que, dentro de dichos límites, las predicciones se pueden considerar completamente acertadas. Este hecho representa un éxito importante dentro de la segunda parte de los objetivos de este trabajo, que consistía en generar modelo ARIMA capaces de predecir con un alto grado de exactitud el comportamiento futuro de las series estudiadas.

5.2.5. Conclusiones sobre las predicciones con modelos ARIMA.

Finalizamos este apartado haciendo un breve resumen de las conclusiones que han ido extrayendo sobre el uso de modelos ARIMA para realizar predicciones en series temporales:

- Tal y como propone la teoría de series temporales, los modelos ARIMA, si han sido bien ajustados, son capaces de predecir con un alto grado de exactitud el comportamiento a futuro de las series temporales.
- En series temporales que se vean afectadas por impactos externos de gran magnitud, las predicciones pueden desviarse en gran medida. Por lo tanto, es necesario actualizar periódicamente los datos de las series temporales y reajustar nuevos modelos ARIMA, con el fin de “perseguir” lo más de cerca posible estos posibles impactos externos. Son precisamente estos impactos los que general los valores atípicos que se detectan en las series temporales, y que por definición no pueden ser predichos si es la primera que ocurren, No obstante, mediante la detección de estos atípicos y su inclusión en el modelo, atípicos iguales o similares si pueden ser predichos, permitiéndonos anticiparnos a ellos,
- Cuando se cuenta con un conjunto grande datos y se realicen predicciones en periodos de relativa estabilidad, las predicciones echas con modelos ARIMA bien ajustados tendrán un alto grado de exactitud.

6 CONCLUSIONES

A raíz de los resultados obtenidos a lo largo de este trabajo, las conclusiones que se pueden extraer son las siguientes:

- Los efectos devastadores de la crisis en varios sectores de la economía española se ven totalmente reflejados en el análisis de series temporales. Durante los años de la crisis, se observa una alta proliferación de outliers (más de 50 atípicos detectados), repartidos de forma homogénea entre todos los sectores estudiados. Se puede afirmar, por lo tanto, que el hecho de incluir la detección de valores atípicos en el análisis de las series ha sido acertado. De no haber sido incluidos estos atípicos en los modelos ARIMA, las predicciones hechas a partir de ellos habrían poseído una precisión notablemente menor, ya que los modelos no serían correctos
- Los modelos ARIMA son una poderosa herramienta para el análisis de series temporales. Mediante el análisis de las series y el estudio de los atípicos detectados, se pudo observar la evolución y desarrollo de la crisis económica en España a través de los sectores estudiados. Recopilando y estudiando el orden de aparición de los atípicos, se puede corroborar como en España la crisis tubo un origen en el sector inmobiliario, provocando un efecto en cadena que arrastró al sector inmobiliario, y acabó afectando a la industria metalúrgica, proveedora de materias primas para el sector inmobiliario. Además, estos valores atípicos no corresponden meramente con fenómenos matemáticos, sino que se ha comprobado que hay acontecimientos reales del panorama social y económico en nuestro país que respaldan su presencia.
- Los modelos ARIMA también han demostrado ser efectivos a la hora de realizar predicciones sobre el comportamiento futuro de las series temporales. Aun así, un modelo matemático nunca podrá ser capaz de predecir el futuro con exactitud, adelantándose a impactos externos fuera del modelo. Siempre y cuando la serie atraviese periodos de relativa estabilidad, las predicciones en base a un modelo ARIMA correctamente estimado tendrán un alto grado de exactitud. Como se ha visto en las estimaciones hechas sobre la serie elegida para tal fin, el grado de exactitud para los datos publicados de forma posterior a la realización de este trabajo tienen un 100% de exactitud dentro de los límites de confianza.

7 ANEXO: CODIGO EN R (V 3.5.1) PARA EL ANÁLISIS DE SERIES TEMPORALES.

A1. codigo para serie temporal “parados en el sector de la construcción”

```
#directorio de trabajo y paquetes#

setwd("C:/Users/Froilan/Desktop/TFG")

install.packages("foreign")
library(foreign)
require(foreign)

install.packages("tseries")
library(tseries)
require(tseries)

install.packages("ggplot2")
library(ggplot2)
require(ggplot2)

install.packages("forecast")
library(forecast)
require(forecast)

install.packages("openxlsx")
library(openxlsx)

install.packages("tsoutliers")

library(tsoutliers)
install.packages("TSA")
library(TSA)

#lectura de datos desde SPSS#
ParadosCONS=read.spss("S22.sav", use.value.labels = TRUE, to.data.frame =
TRUE)

#definición como serie temporal#
```

```

vParadosCONS= ParadosCONS[,1]
tsParadosCONS=ts(vParadosCONS, start=c(1996, 1), end=c(2017, 9),
frequency=12)

#gráfico de secuencia y grafico de series aditivas#

plot(tsParadosCONS)
plot(decompose(tsParadosCONS))

#Gráfico TSO para ARIMA ajustado en SPSS#
tso(y = log(tsParadosCONS), cval = 3.5,
types = c("AO", "LS", "TC", "SLS", "IO"), maxit = 1,
remove.method = "bottom-up", tsmethod = "arima",
args.tsmethod=list(order=c(1,1,1),
seasonal = list(order = c(2, 1, 0))))

```

A2. código para serie temporal “parados en el sector de la construcción”

```

#lectura de datos desde SPSS#
IndPRECMO=read.spss("S2error.sav", use.value.labels = TRUE, to.data.frame =
TRUE)

#definición como serie temporal#
vIndPRECMO=IndPRECMO[,1]
tsIndPRECMO=ts(vIndPRECMO, start=c(2002, 1), end=c(2016, 12),
frequency=12)

#gráfico de secuencia y transformación en estacionaria (tsdisplay)#
tsdisplay(tsIndPRECMO)
tsdisplay(diff(diff(tsIndPRECMO), lag=12, differences=1))

#auto ARIMA como comprobación#
auto.arima(tsIndPRECMO)

#TSO y gráfico TSO#
tso(y = tsIndPRECMO, cval = 3.5,types = c("AO", "LS", "TC", "SLS", "IO"),maxit
= 1, remove.method = "bottom-up", tsmethod = "arima", args.tsmethod = list(order
= c(1, 1, 0), seasonal = list(order = c(0, 1, 1))))

plot(tso(y = tsIndPRECMO, cval = 3.5,types = c("AO", "LS", "TC", "SLS",
"IO"),maxit = 1, remove.method = "bottom-up", tsmethod = "arima", args.tsmethod
= list(order = c(1, 1, 0), seasonal = list(order = c(0, 1, 1))))

```

```
#comprobación y estudio de los residuos a partir de TSO#
residualstsIndPRECMO=residuals(TSOIndPRECMO$fit)
acf(residualstsIndPRECMO)
Pacf(residualstsIndPRECMO)
Box.test( residuals(TSOIndPRECMO$fit), type = c("Box-Pierce"))
Box.test( residuals(TSOIndPRECMO$fit), type = c("Ljung-Box"))
acf(residualstsIndPRECMO)
plot(residualstsIndPRECMO)
```

A3. código para la serie “asociados a la seguridad social en el sector de la construcción”.

```
#lectura de datos desde SPSS#
SSocialCONS=read.spss("S25.sav", use.value.labels = TRUE, to.data.frame =
TRUE)

#definición como serie temporal#
vSSocialCONS=IndPRECMO[,1]
tsSSocialCONS=ts(vSSocialCONS, start=c(1995, 1), end=c(2017, 9),
frequency=12)

#gráfico de secuencia y transformación en estacionaria (tsdisplay)#
tsdisplay(tsSSocialCONS)
tsdisplay(diff(diff(tsSSocialCONS), lag=12, differences=1))

#test aumentado de Dickey-Fuller#
adf.test(diff(diff(tsSSocialCONS), lag=12, differences=1), alternative=
c("stationary"))

#TSO y gráfico TSO#
tso(y = tsSSocialCONS, cval = 3.5,types = c("AO", "LS", "TC", "SLS", "IO"),maxit
= 1, remove.method = "bottom-up", tsmethod = "arima", args.tsmethod = list(order
= c(1, 1, 0), seasonal = list(order = c(0, 1, 1))))

plot(tso(y = tsSSocialCONS, cval = 3.5,types = c("AO", "LS", "TC", "SLS",
"IO"),maxit = 1, remove.method = "bottom-up", tsmethod = "arima", args.tsmethod
= list(order = c(1, 1, 0), seasonal = list(order = c(0, 1, 1))))

#comprobación y estudio de los residuos a partir de TSO#
residualstsSSocialCONS =residuals(TSOtsSSocialCONS$fit)
acf(residualstsSSocialCONS)
Pacf(residualstsSSocialCONS)
Box.test( residualstsSSocialCONS, type = c("Box-Pierce"))
Box.test( residualstsSSocialCONS), type = c("Ljung-Box"))
Tsdisplay(residualstsSSocialCONS)
```


Para el resto de series del trabajo el código utilizado es el mismo, simplemente cambiando el nombre por el de cada una de las series.

8 BIBLIOGRAFÍA

- [1] D. Peña, *Análisis de series temporales*. Madrid: Alianza Editorial, 2017.
- [2] G. Box and G. Jenkins, *Time series analysis*. Oakland: Holden-Day, 1976.
- [3] J. López de Lacalle, *Detection of Outliers in Time Series*, 2017. [Online]. Disponible en: <https://cran.r-project.org/web/packages/tsoutliers/tsoutliers.pdf>.
- [4] H. Aznar, "Evolución en España del Sector de la Metalurgia No Ferrea", *TECNOINDUSTRIA*, 2016. [Online]. Disponible en: <https://technoindustria.wordpress.com/2016/02/16/evolucion-en-espana-del-sector-de-la-metalurgia-no-ferrea/>.
- [5]: IBM (2012) Manual del usuario del sistema del sistema básico de IBM SPSS Statistics 23.
- [6] Euribor, 2018. [Online]. Disponible en: <http://www.hipotecasyeuribor.com/mes-Marzo2008.htm>.
- [6] "La construcción cayó en febrero el 5,3% en España frente al 1,8% en la eurozona", *Levante-emv.com*, 2009. [Online]. Disponible en: <https://www.levante-emv.com/economia/2009/04/17/construccion-cayo-febrero-53-espana-frente-18-eurozona/579307.html>.
- [7] "Las fechas decisivas en la crisis económica española", *Elmundo.es*, 2012. [Online]. Disponible en: <http://www.elmundo.es/elmundo/2012/06/09/economia/1339243054.html>.
- [8] C. Pérez, "La crisis en España: cronología desde 2008 - RTVE.es", *RTVE.es*, 2012. [Online]. Available: <http://www.rtve.es/noticias/20120605/crisis-espana-cronologia-desde-2008/533400.shtml>.
- [9] J. Gonzalez Cuesta, "EL CRASH DE 2008 | Economía | Especiales | elmundo.es", *Elmundo.es*, 2018. [Online]. Available: <http://www.elmundo.es/especiales/2008/10/economia/crisis2008/espana/index.html>.
- [10] S. López Letón, "La burbuja que embriagó a España", *EL PAÍS*, 2015. [Online]. Available: https://elpais.com/economia/2015/10/20/actualidad/1445359564_057964.html.

[11] M. Fortuño, "Así han sido los últimos diez años del sector inmobiliario en España", Elblogsalmon.com, 2017. [Online]. Available: <https://www.elblogsalmon.com/sectores/asi-han-sido-los-ultimos-diez-anos-del-sector-inmobiliario-en-espana>.

[12] R. Tamames and A. Rueda, *Estructura económica de España*. 2008.